Nachrichten

von der Königlichen Gesellschaft der Wissenschaften zu Göttingen

Mathematisch-physikalische Klasse
aus dem Jahre 1919

BERLIN
Weidmannsche Buchhandlung
1919

Druck der Dieterichschen Univ.-Buchdruckerei (W. Fr. Kaestner) in Göttingen.

Register

über

die Nachrichten von der Königl. Gesellschaft der Wissenschaften zu Göttingen.

Mathematisch-physikalische Klasse

aus dem Jahre 1919.

Über die Strömungspotentiale und die zugehörenden konformen Abbildungen Riemannscher Flächen.

Von

Herrn **Paul Koebe** in Jena.

Vorgelegt in der Sitzung vom 13. Dezember 1918.

Wird mit U diejenige in einem schlichten oder schlichtartigen q-fach zusammenhängenden Bereich B eindeutige Potentialfunktion bezeichnet, welche nur in einem Punkte innerhalb B unstetig wird und zwar wie $r^{-1}\cos\varphi$ im Punkte $r = 0$, außerdem längs der ganzen Begrenzung die normale Ableitung null hat, so vermittelt die analytische Funktion $U + iV = W$ bekanntlich eine eineindeutige konforme Abbildung des Bereiches B auf einen schlichten Schlitzbereich, der von der ganzen längs q endlichen Strecken parallel der Achse des Reellen aufgeschlitzten W-Ebene gebildet wird. Dieser Satz wird im folgenden Verallgemeinerungen unterworfen, zu denen ich bereits 1909 in Verfolgung gewisser damals von Herrn Hilbert in einer Note in den Gött. Nachr.[1]) aufgestellter Behauptungen gelangt bin.

Wir betrachten eine beliebige geschlossene oder berandete Riemannsche Fläche F von endlichem Zusammenhange und auf derselben alle möglichen eindeutigen Potentiale U, deren normale Ableitung längs der ganzen Berandung gleich null ist, d. s. alle

---

1) D. Hilbert: „Zur Theorie der konformen Abbildung“, Gött. Nachr. 1909. Wegen Abbildung unendlich-vielfach zusammenhängender schlichtartiger Bereiche auf schlichte Schlitzbereiche s. meine Abhandlungen: „Über die Uniformisierung beliebiger analytischer Kurven“ in Gött. Nachr. 1909 und Journal f. Math. Bd. 138; „Über die Hilbertsche Uniformisierungsmethode“ in Gött. Nachr. 1910; „Zur konformen Abbildung unendlich vielfach zusammenhängender schlichter Bereiche auf Schlitzbereiche“ in Gött. Nachr. 1918. S. auch dort genannte Arbeiten von R. Courant. S. auch Fußnote am Ende.

diejenigen Strömungspotentiale, die, von inneren sowie auch Randquellpunkten in beliebiger Anzahl und von beliebiger Ordnung herrührend, unter der Voraussetzung völliger Isolierung längs des ganzen Randes zustandekommen. Die zugehörenden analytischen Funktionen sind charakterisierbar als die allgemeinsten Abelschen Integrale mit rein imaginären Periodizitätsmoduln und längs jedes quellenfreien Teiles der Begrenzung konstantem imaginären Teil.

Bei einer einfachen Dipolströmung auf einer geschlossenen Riemannschen Fläche F (§§ 1 u. 2) ergibt sich eine Abbildung dieser in gewisser Weise längs „singulären Stromlinien" aufgeschnitten zu denkenden Fläche auf einen schlichten Bereich $\tilde{\Sigma}_1$, der die W-Ebene überall einfach bedeckt und im allgemeinen von endlich vielen ($4p$, wenn p das Geschlecht von F ist) nach der Seite des negativ Reellen ins Unendliche gehenden paarweise auf einander bezogenen Schlitzen parallel der Achse des Reellen begrenzt wird. Ist die Fläche F berandet (§ 3) und bedeutet dann p die maximale Anzahl nicht zerstückender innerer Rückkehrschnitte, die man ziehen kann, q die Anzahl der Randlinien, so wird der geschilderte Bereiche $\tilde{\Sigma}_1$ im allgemeinen durch das Hinzukommen von q endlichen, von erstgenannten Begrenzungsschlitzen ganz getrennt liegenden unbezogenen Begrenzungsschlitzen modifiziert (Bereich $\tilde{\Sigma}_2$). Wird der Dipol durch zwei getrennte Punkte (Quellpunkt und Senkpunkt (logarithmische Unstetigkeiten)) ersetzt (§ 3), so ergibt sich ein Bereich $\tilde{\Sigma}_3$ in Gestalt eines aufgeschlitzten Parallelstreifens endlicher Breite bzw. bei Übergang zur Funktion e^{U+iV} wieder eine Vollebene, die im allgemeinen von $4p$ vom Nullpunkte ausgehenden paarweise auf einander bezogenen geradlinigen endlichen Schlitzen begrenzt wird. Werden schließlich Unstetigkeiten in beliebiger Anzahl und von beliebiger Ordnung im Innern sowie auch am Rande zugelassen (§ 4), so ergibt sich, allgemein gesprochen, als Bild der aufgeschnittenen Fläche F ein Bereich $\tilde{\Sigma}$, der durch Übereinanderheftung von endlich vielen parallel der Achse des Reellen aufgeschlitzten Blättern (Vollblättern, Halbblättern, streifenförmigen Blättern) entsteht.

Die gewonnenen Abbildungsresultate eröffnen die Möglichkeit, auch über die Natur der betrachteten Strömungen nähere Aufschlüsse zu erlangen. So wird bei der einfachen Dipolströmung auf geschlossenen Riemannschen Flächen vom Geschlecht p gezeigt, daß die Anzahl λ der wesentlich verschiedenen nicht weiter ge-

spaltenen „Stromverbände“ („Stromgebiete“) jeden Wert von 3 bis $4p+1$ annehmen kann, sofern p von null verschieden ist; ferner, daß die Anzahl N getrennt liegender Staupunkte, deren Gesamtordnung gleich $2p$ sein muß, jeden Wert von 1 bis $2p$ annehmen kann und daß auch eine beliebige Verteilung der Gesamtordnung auf die einzelnen Staupunkte erzielt werden kann. Analoge Resultate ergeben sich auch, wenn die Fläche F berandet ist; den Randstaupunkten ist dabei nur halbes Gewicht beizumessen.

Inhaltsverzeichnis.

Erster Teil.

Einfache Dipolströmung und zugehörende Abbildung geschlossener Riemannscher Flächen.

1. Einführung der Funktion $U+iV$. Es werde mit F eine geschlossene Riemannsche Fläche vom Geschlecht p bezeichnet, mit U das in F eindeutige Strömungspotential, das in einem Punkte O der Fläche F unstetig wird wie $r^{-1}\cos\varphi$. Der Punkt O werde dabei als ein gewöhnlicher Punkt der Fläche F gewählt. Ist O ein Windungspunkt m-ter Ordnung der Fläche F, so ist die Unstetigkeit $r^{-1}\cos\varphi$ durch die andere $r^{-\frac{1}{m+1}}\cos\frac{\varphi}{m+1}$ zu ersetzen, die durch eine den Windungspunkt in einen gewöhnlichen Punkt verwandelnde Wurzeltransformation in erstere übergeht. Ist O ein unendlich ferner Punkt auf F, so hätte man als Unstetigkeit $r\cos\varphi$ bezw. $r^{\frac{1}{m+1}}\cos\frac{\varphi}{m+1}$ zu wählen entsprechend

den diesen Punkt der Fläche in einen gewöhnlichen endlichen Punkt verwandelnden lokalen Transformationen. Weiter steht es frei, die Unstetigkeit noch zu erweitern zu der nicht wesentlich allgemeineren Unstetigkeit

$$ar^{-1}\cos\varphi + br^{-1}\sin\varphi = \sqrt{a^2+b^2}\cos(\varphi-\varphi_0),$$

was physikalisch nur eine Änderung der Stärke des Dipols und eine Richtungsänderung der Dipolachse, geometrisch eine Drehung des Koordinatensystems in Verbindung mit einer Änderung des Maßstabes bedeutet.

Ohne Beschränkung des Allgemeinheitsgrades unserer Untersuchung können wir uns auf die Betrachtung des Falles eines gewöhnlichen Punktes O und der Unstetigkeit $r^{-1}\cos\varphi$ beschränken. Die Funktion U sei irgendwie gewonnen, z. B. potentialtheoretisch mit Hilfe des alternierenden Verfahrens oder auch aus dem zur Fläche F gehörenden algebraischen Funktionenkörper, dessen Existenz auch rein funktionentheoretisch begründet werden kann [1]). Es werde mit V die zu U gehörende konjugierte Potentialfunktion bezeichnet. Diese wird nur in O unstetig, nämlich wie $-r^{-1}\sin\varphi$; sie wird aber im Gegensatz zu U im allgemeinen auf F nicht eindeutig sein. Die Funktion $U+iV$ ist ein Abelsches Integral zweiter Art mit rein imaginären Periodizitätsmoduln und nur einer Unendlichkeitsstelle erster Ordnung im Punkte O ($z = z_0$). Man hat dort eine Entwicklung der Form

$$U+iV = \frac{1}{z-z_0} + \text{reg. Funkt.}.$$

Eindeutigkeit der Funktion $U+iV$ kann, wie sich auch im weiteren Verlaufe der Untersuchung von selbst ergeben wird, nur eintreten, wenn das Geschlecht p der Fläche F gleich null ist. Denn im Falle der Eindeutigkeit vermittelt die Funktion eine eineindeutige Abbildung der Fläche F auf eine Riemannsche Fläche, die die Ebene einerseits überall gleich-vielblättrig, andererseits den unendlich fernen Punkt jedenfalls nur einfach bedeckt, womit gesagt ist, daß letztere Riemannsche Fläche mit der schlichten Ebene identisch, die Fläche F selbst demgemäß vom Geschlecht null sein muß.

2. Allgemeiner Stromlinienverlauf. Wir lenken jetzt unsere Aufmerksamkeit zunächst auf den allgemeinen Verlauf der

1) S. meine Abhandlung II der Serie „Abhandlungen zur Theorie der konformen Abbildung“ in Acta math. Bd. 40, insbesondere § 8.

Linien $V = \text{const.}$, d. i. der Stromlinien. Die Stromrichtung identifizieren wir mit derjenigen, in welcher die Potentialwerte U abnehmen. Der Verlauf der Stromlinien im Quellpunkte O wird durch Fig. 1 dargestellt[1]). Wir bemerken noch besonders, daß ein Teil der Stromlinien wie in Fig. 1 angedeutet ist, vollständig in unmittelbarer Nähe des Quellpunktes verlaufen und also in der Figur als geschlossene Linien erscheinen.


Figur 1.

Die Funktion $U + iV$ wird eine gewisse konforme Abbildung der Riemannschen Fläche F vermitteln, wobei die Umgebung eines von O verschiedenen Punktes entweder, und zwar im allgemeinen, auf ein schlichtes endliches Gebiet abgebildet wird („gewöhnlicher Punkt") oder aber auf ein Windungsflächenstück (Staupunkt). Letzteres tritt nur ein an den Nullstellen des Differentials $dW = dU + idV$. Die Anzahl dieser Nullstellen ist bei Berücksichtigung ihrer Vielfachheit gleich $2p$ nach dem Satze, daß für jedes Abelsche Differential die Anzahl seiner Nullstellen vermindert um die Anzahl seiner Unendlichkeitsstellen gleich $2p-2$ ist[2]). Für das Folgende genügt es, jetzt die Endlichkeit der Anzahl der Staupunkte festzustellen. Diese aber folgt einfach daraus, daß bei Annahme des Gegenteils ein Häufungspunkt von Nullstellen des Differentiales vorhanden sein würde, was unmöglich ist. Die genaue Zahl $2p$ wird sich weiter unten von selbst ergeben.

In den „gewöhnlichen Punkten" zeigt die Strömung den in Fig. 2 dargestellten Verlauf: ein glattes Nebeneinanderherlaufen der Stromlinien. Windungspunkte und unendlich ferne Punkte (als „gewöhnliche Punkte") hat man sich dabei stillschweigend der


Figur 2. Figur 3 a. Figur 3 b.

üblichen lokalen Hilfstransformation unterworfen zu denken. Der Verlauf der Stromlinien an den Staupunkten wird durch die Fi-

1) Wegen des Verlaufs der Stromlinien an Strömungspolen und Staupunkten vgl. man unsere Figuren 1, 3, 25 mit den bezüglichen Figuren in Kleins Schrift „Über Riemanns Theorie usw." (Teubner).

2) Vgl. S. 9, 10.

guren 3 dargestellt, deren Schema sich ergibt, wenn man in ein Windungsflächenstück die Linien parallel der Achse des Reellen einzeichnet und sodann dieses Liniensystem derjenigen Wurzeltransformation unterwirft, die das Windungsflächenstück in ein schlichtes Gebiet verwandelt. Diese Figuren erfahren wieder eine sinngemäße Modifikation, wenn die Staupunkte mit Windungspunkten oder unendlich fernen Punkten der Riemannschen Fläche zusammenfallen.

Wir verfolgen nunmehr, von einem „gewöhnlichen Punkte" der Fläche F ausgehend, den Verlauf der durch ihn gehenden Stromlinie nach beiden Seiten, soweit als sie im Bereiche „gewöhnlicher Punkte" bleibt. Wir behaupten, daß wir dabei schließlich notwendig entweder zum Quellpunkte O oder zu einem Staupunkte gelangen, ohne daß die Linie in ihrem Verlaufe sich selbst begegnen kann. Eine Selbstbegegnung bezw. ein Insichzurücklaufen der Linie ist nämlich deswegen unmöglich, weil bei Durchlaufung der Linie die Größe U sich monoton ändert und folglich die erwähnten Vorkommnisse einen Widerspruch mit der Eindeutigkeit der Funktion U liefern würden.

Würde nun die Stromlinie weder zum Quellpunkte noch zu einem Staupunkte führen, so würde sie dem Quellpunkte auch nicht beliebig nahe kommen können, da sie ja andernfalls wegen des Stromlinienverlaufes am Quellpunkte (Fig. 1) notwendig in den Quellpunkt selbst führen müßte. Es würde sich infolgedessen ein vom Quellqunkte verschiedener, ev. mit einem Staupunkte zusammenfallender Punkt auf der Fläche F ergeben, in dessen beliebiger Nähe die betreffende Stromlinie noch beliebig oft vorbeiströmt. Aus dem lokalen Stromlinienverlaufe an diesem Häufungspunkte ergibt sich dann, daß die in einer gewissen Umgebung dieses Punktes bei Verfolgung der Stromlinie zustande kommenden Zuwüchse von U beim jedesmaligem Passieren dieser Umgebung eine gewisse von null verschiedene untere Schranke übertreffen, sodaß die Größe U in der Nachbarschaft des Häufungspunktes beliebig große Werte erlangen würde, was unmöglich ist.

3. Stromverzweigungsnetz, Stromgebiete (Stromverbände). In jedem Staupunkt treffen eine endliche Anzahl Stromlinien zusammen, nämlich $2\tau + 2$, wenn τ die Ordnung der Vielfachheit des Staupunktes ist. Diese $2\tau + 2$ Linien bilden im Staupunkte $2\tau + 2$ unter einander gleiche Winkel der Größe $\frac{\pi}{\tau + 1}$ (vgl. Fig. 3) bzw. $\frac{(m+1)\pi}{\tau + 1}$, wenn es sich um einen Windungspunkt m-ter Ordnung handelt. Indem wir diese Linien von den

Staupunkten aus verfolgen, werden wir bei der einzelnen Linie entweder zu einem anderen Staupunkte oder zum Quellpunkte gelangen. Das so gewonnene Liniensystem werde als das Stromverzweigungsnetz bezeichnet; die einzelnen Linien desselben mögen singuläre Stromlinien heißen. Durch das Stromverzweigungsnetz wird die Fläche F in eine endliche Anzahl λ von Gebieten (Stromgebiete) zerlegt. Indem wir den Rand eines Stromgebietes abschreiten, kann U nur eine monotone Änderung erfahren. Daraus ergibt sich, daß der Rand eines Stromgebietes stets den Quellpunkt O enthalten muß. Die Anzahl der in O beginnenden singulären Stromlinien muß gleich sein der Anzahl der nach in O mündenden singulären Stromlinien. Dies folgt daraus, daß man bei Abschreitung des Randes eines Stromgebietes, ausgehend von O, wegen der beständigen monotonen Änderung von U notwendig von der anderen Seite her wieder nach null gelangen muß. Es besteht demnach eine ganz bestimmte Uferpaarung der in O beginnenden und in O endigenden singulären Stromlinien.

Gehen wir von irgend einem inneren Punkte eines Stromgebietes, der kein Randpunkt desselben ist, aus, so muß nach der Definition des Stromverzweigungsnetzes die durch diesen Punkt hindurchgehende Stromlinie eine einfache Verbindung der positiven Seite des Quellpunktes (U positiv, Einströmungsseite) mit der negativen Seite des Quellpunktes (U negativ, Abströmungsseite) herstellen, ohne sich selbst zu begegnen oder auf den Rand des betreffenden Stromgebietes zu treffen. Wir können uns daher eine vollständige Übersicht über den ganzen Stromlinienverlauf verschaffen, wenn wir in O alle Stromlinienanfänge nehmen und die zugehörenden vollständigen Stromlinien konstruieren wobei durch die singulären Stromlinienanfänge eine natürliche Einteilung gegeben ist.

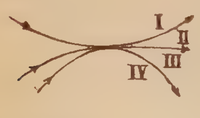


Fig. 4.

Nehmen wir um die Vorstellung zu fixieren, etwa den in Fig. 4 schematisch dargestellten Fall mit drei singulären Einströmungslinien an. Indem wir die Stromlinienanfänge in den Bereichen I, II, III, IV wählen, müssen wir mit den zugehörenden vollständigen Stromlinien jedesmal ein vollständiges Stromgebiet überstreichen. Aus der einfachen glatten Nebeneinanderlagerung der Stromlinien in jedem dieser vier Stromgebiete ziehen wir den Schluß, daß jedes Stromgebiet einfach zusammenhängend ist. Die Stromlinien eines solchen Gebietes bilden in ihrer Gesamtheit einen Stromverband. Zwei verschiedene Stromverbände können nach unserer Begriffsbestim-

mung, sofern sie nicht völlig getrennt verlaufen, nur teilweise, evtl. zu wiederholten Malen nebeneinander herlaufen. Wir bemerken schließlich, daß an den beiden äußersten Stromgebieten (I und IV) der Punkt O nur einmal als Randpunkt teilnimmt, während er bei den mittleren Stromgebieten (II. III.) zweimal als Randpunkt teilnimmt, indem jeder der mittleren Stromverbände in O zwei Enden hat, ein Einströmungs- und ein Abströmungsende. Die in Rede stehenden vier einfach zusammenhängenden Stromgebiete bedecken notwendig die ganze Fläche F. Es sind diese Gebiete eben nichts anderes, als die durch das Stromverteilungsnetz definierten Stromgebiete, deren Anzahl wir allgemein mit λ bezeichneten. Die Zahl $\lambda - 1$ hat somit die zweite Bedeutung als Anzahl der in O beginnenden singulären Stromlinien, eine Zahl, die übereinstimmt mit der Anzahl der in O mündenden singulären Stromlinien.

4. **Anzahl der Staupunkte.** Wie wir bereits oben erwähnten, ist die Anzahl der Staupunkte bei Berücksichtigung ihrer Vielfachheit gleich $2p$. Wir haben von dieser Bemerkung, die nur die Anwendung eines allgemeinen Satzes über Abelsche Differentiale auf unser Differential dW darstellt, bisher keinen Gebrauch zu machen gehabt. Nunmehr wollen wir uns überzeugen, daß dieser Satz sich als eine Folge aus den vorstehenden Entwicklungen ergibt.

Zu dem Zwecke bezeichnen wir mit σ die zu ermittelnde Anzahl der Staupunkte bei Berücksichtigung ihrer Vielfachheit, mit N die Anzahl der voneinander wirklich verschiedenen Staupunkte, mit s die Anzahl der singulären Stromlinien, deren einzelne für uns immer eine von Staupunkt zu Staupunkt oder von Staupunkt zu Quellpunkt führende Stromlinie sei, bei deren vollständiger Durchlaufung zwischendurch kein Staupunkt angetroffen wird. Mit λ wurde allgemein die Anzahl der Stromgebiete bezeichnet, in die die Fläche F durch das von den singulären Stromlinien gebildete Stromverzweigungsnetz zerlegt wird. Indem wir die Fläche F im Punkte O und in sämtlichen Staupunkten punktieren, entsteht aus F eine Fläche f vom Zusammenhange $2p + N + 1$. Bezeichnet τ_α die Ordnung der Vielfachheit des α-ten Staupunktes, so treffen in diesem Punkte $2(\tau_\alpha + 1)$ singuläre Stromlinien zusammen. Im Punkte O treffen $2(\lambda - 1)$ singuläre Stromlinien zusammen. Da bei dieser Berechnung jede singuläre Stromlinie zweimal gezählt wird, so ist die Anzahl s der singulären Stromlinien gleich

$$s = \sum_{\alpha=1}^{N} (\sigma_\alpha + 1) + (\lambda - 1) = \sum_{\alpha=1}^{N} \sigma_\alpha + N + \lambda - 1 = \sigma + N + \lambda - 1.$$

Die Fläche f wird durch diese s Linien in λ einfach zusammenhängende Gebiete zerlegt, daraus folgt die Gleichung

$$2p + N + 1 - (\sigma + N + \lambda - 1) = -(\lambda - 2)$$

d. h.

$$\sigma = 2p.$$

5. Abbildung der Fläche F auf einen schlichten Schlitzbereich $\tilde{\Sigma}_1$. Wir betrachten zunächst den Fall des gänzlichen Fehlens von Staupunkten. In diesem Falle haben wir keine singulären Stromlinien, mithin ein einziges Stromgebiet, einen einzigen Stromverband. Sämtliche in O beginnenden Stromlinien lagern sich in ihrem vollständigen Verlaufe glatt nebeneinander und münden notwendig in gleicher Ordnung von der anderen Seite her wieder in O die einzelnen Stromlinien können dabei durch den auf ihr bestehenden V-Wert definiert werden. Dieser Wert variiert von $+\infty$ bis $-\infty$. Eine unmittelbare Folge des gefundenen Stromlinienverlaufes ist die Eindeutigkeit der Abbildung der vollständigen Fläche F durch die Funktion $U + iV$ auf die volle schlichte Ebene, auf welche übertragen die Stromlinien mit den Geraden parallel der Achse des Reellen identisch werden, wobei die übertragene Strömung als eine von rechts aus dem Unendlichen kommende aufzufassen ist.

In dem allgemeinen Falle, $p > 0$, sind $2p > 0$ Staupunkte vorhanden, daher auch ein Stromverzweigungsnetz und somit auch eine Anzahle $\lambda > 1$ von Stromgebieten, die sich später immer als $\geqq 3$ erweisen wird. Wir gehen von dem System der λ einfach zusammenhängenden Stromgebiete zu einem einfach zusammenhängendem Bereiche $\overline{F}$ dadurch über, daß wir zwischen je zwei bei O auf der Einströmungsseite benachbart liegenden Stromgebieten, eine Verbindung herstellen. Der einfachste Modus ist der, daß wir die in O beginnenden singulären Stromlinien beseitigen. In der so entstehenden einfach zusammenhängenden Fläche $\overline{F}$ ist die Funktion $U + iV$ notwendig eindeutig bis auf die längs der Berandung von $\overline{F}$ sich einstellenden rein imaginären Periodizitätsmoduln. Durch die Funktion $U + iV$ werden die beiden äußersten Stromgebiete, in denen V beliebige große positive bezw. beliebig große negative Werte annimmt, wegen des bekannten Stromlinienverlaufs offenbar je auf eine Halbebene abgebildet, während die $\lambda - 2$ mittleren Stromgebiete je auf einen Parallelstreifen abgebildet werden. Diese zwei Halbebenen und $\lambda - 2$ Parallelstreifen

schließen sich entsprechend der zwischen den λ Stromgebieten bei O hergestellten Verbindung zu einem einfach zusammenhängenden Bereiche $\overline{\Sigma}$ zusammen, der von der ganzen schlichten Ebene gebildet wird, die dabei durch $\lambda - 1$ aus dem Endlichen kommende nach der Seite des Negativen ins Unendliche gehende Schlitze aufgeschnitten erscheint. Diese Schlitze bilden die vollständige Berandung des Bereiches $\overline{\Sigma}$. Die endlichen Anfangspunkte derselben entsprechen Staupunkten der Fläche F. Wir markieren auf diesen Schlitzen, deren je zwei Ufer (oberes und unteres Ufer) voneinander unterschieden werden müssen, diejenigen weiteren Punkte, die Staupunkten der Fläche F entsprechen. Die so gewonnene Einteilung der Berandung von $\overline{\Sigma}$ in Stücke ist von der Art, daß je zwei dieser Stücke, wovon immer eins einem oberen Ufer, das andere einem unteren Ufer angehört, einander durch eine Parallelverschiebung parallel der Achse des Imaginären entsprechen. Hierbei ist auch der Fall des Entsprechens durch Koinzidenz einbegriffen. Die rechts liegenden Endpunkte solcher durch Koinzidenz aufeinander bezogenen Stücke können jedoch niemals mit einem Endpunkte unserer $\lambda - 1$ Begrenzungsschlitze des Bereiches $\overline{\Sigma}$ zusammenfallen; wohl aber können diese Stücke sich nach links bis ins Unendliche erstrecken.

Indem wir alle durch Koinzidenz aufeinander bezogenen Randteile des Bereiches $\overline{\Sigma}$ entfernen, entsteht aus $\overline{\Sigma}$ ein neuer, allgemein zu reden mehrfach zusammenhängender, ebenfalls schlichter Schlitzbereich $\tilde{\Sigma}_1$ mit bezogenen Rändern. Dieser Bereich $\tilde{\Sigma}_1$ ist das eindeutige Abbild einer Fläche $\tilde{F}$, die aus $\overline{F}$ durch Fortlassung aller derjenigen Randstücke dieses Bereiches entsteht, längs welchen keine Peridizitätsmoduln vorhanden sind, physikalisch gesprochen, längs welchen ursprünglich bei O benachbarte Stromverbände in Wiedervereinigung erscheinen.

Zweiter Teil.

Nähere Diskussion und Charakterisierung der Schlitzbereiche $\tilde{\Sigma}_1$.

6. Einige allgemeine Bemerkungen über die Natur der Schlitzbereiche $\tilde{\Sigma}_1$. Der als Bild der Fläche $\tilde{F}$ durch Vermittlung der Funktion $U + iV$ gewonnene Schlitzbereich $\tilde{\Sigma}_1$ soll nunmehr selbständig untersucht werden. Dabei werden sich eine ganze Reihe möglicher Fälle ergeben.

Wir erinnern zunächst daran, daß der Bereich $\bar{\Sigma}$ aus $\tilde{\Sigma}_1$ entsteht, indem man sämtliche Begrenzungsschlitze von $\tilde{\Sigma}_1$ nach links bis ins Unendliche verlängert, wobei es allgemein zu reden vorkommt, daß eine solche Verlängerung mehrere getrennte Begrenzungsschlitze des Bereichs $\tilde{\Sigma}_1$ enthält. Die Zahl der so entstehenden Schlitze ist, der früheren Bedeutung von λ entsprechend, mit $\lambda-1$ zu bezeichnen. Diese Zahl $\lambda-1$ muß mindestens gleich 2 sein, d. h. für die Strömung auf F: Die Zahl λ der Stromverbände muß bei $p>0$ immer $\geqq 3$ sein. Andernfalls würde nämlich als Bezugssubstitution nur die identische Substitution in Frage kommen. Dies würde aber bedeuten, daß der Bereich $\tilde{\Sigma}_1$ überhaupt keine Begrenzungsschlitze hat, sondern von der unbegrenzten schlichten Ebene gebildet wird. Die schlichte volle W-Ebene würde dann das eineindeutige Abbild der geschlossenen Riemannschen Fläche F sein. Letztere müßte demnach vom Geschlecht null sein, ein Fall, den wir nunmehr von der Betrachtung ausschließen.

Eine weitere allgemeine Bemerkung ist die, daß die $\lambda-1$ unendlich fernen Ecken des Bereiches $\bar{\Sigma}$, von denen $\lambda-2$ den Öffnungswinkel 0, eine den Öffnungswinkel 2π hat, vermöge der bestehenden Ränderzuordnung zu einem Zykel zusammengefaßt sein müssen, was unmittelbar daraus folgt, daß diese $\lambda-1$ Ecken im Original auf F die volle Umgebung des Punktes 0 bilden. Was von den $\lambda-1$ unendlich fernen Ecken des Bereichs $\bar{\Sigma}$ gilt, gilt dann natürlich auch von den unendlich fernen Ecken des Bereichs $\tilde{\Sigma}_1$: auch sie, deren Anzahl $\leqq \lambda-1$ ist, bilden einen Zykel.

Drittens bemerken wir, daß die Zahl λ nicht größer als $4p+1$ sein kann. Vermöge der bestehenden Ränderzuordnung setzen sich nämlich die Umgebungen der rechten Endpunkte unserer $\lambda-1$ Begrenzungsschlitze von $\bar{\Sigma}$ im Sinne der Zykelnbildung zu Windungsflächenstücken zusammen. Wenn hierbei zwei jener Endpunkte einen Windungspunkt liefern, so ist derselbe von der ersten Ordnung. Liefern aber erst $\varkappa$ jener Endpunkte einen Windungspunkt, so ist derselbe von der Ordnung $\varkappa-1$. Es kann schließlich noch der Fall vorliegen, daß einer oder mehrere jener Endpunkte im Sinne der Zykelnbildung mit Teilpunkten auf den Schlitzufern zusammengehören. In diesem Falle ergibt sich, wenn $\varkappa$ die Anzahl der an einem solchen Zykel beteiligten Endpunkte ist, ein Windungspunkt von mindestens $\varkappa$-ter Ordnung. In jedem

Falle zeigt sich, daß die $\lambda - 1$ Endpunkte mindestens $\left[\frac{\lambda-1}{2}\right]$ Windungspunkte liefern bei Berücksichtigung der Vielfachheit der letzteren; dabei bedeutet $\left[\frac{\lambda-1}{2}\right]$ die größte in $\frac{\lambda-1}{2}$ enthaltene ganze Zahl. Da nun die Gesamtzahl der aus dem Bereich $\tilde{\Sigma}_1$ durch Zykelnbildung entstehenden Windungspunkte gleich $2p$ sein muß, nämlich gleich der Zahl der Staupunkte auf F, so kann λ *nicht größer als* $4p+1$ sein.

Eine ganz analoge Betrachtung lehrt viertens, daß der Zusammenhang des Bereiches $\tilde{\Sigma}_1$ nicht größer als $2p$ sein kann. Man bedenke nämlich, daß jeder die Zusammenhangzahl erhöhende vollständige Begrenzungsschlitz von $\tilde{\Sigma}_1$ ein endlicher Begrenzungsschlitz sein muß, der somit einen rechten und einen linken Endpunkt hat. Wäre nun der Zusammenhang größer als $2p$, so würden jene Endpunkte gemäß dem Prinzip der Zykelnbildung offenbar zu mehr als $2p$ Windungspunkten Veranlassung geben.

7. *Natur des Schlitzbereiches* $\tilde{\Sigma}_1$ *in gewissen extremalen Fällen.* Die vorstehenden Betrachtungen geben auch über die genauere Natur der Bereiche $\tilde{\Sigma}_1$ in den erwähnten extremalen Fällen Aufschluß.

Im Falle $\lambda = 3$ finden wir, da als Bezugsubstitutionen des Bereiches $\tilde{\Sigma}_1$ dann lediglich ein und dieselbe Parallelverschiebung in Betracht kommt, folgendes: Der Bereich $\tilde{\Sigma}_1$ wird begrenzt durch p *in gleicher Höhe liegende Paare von je zwei gleich langen genau übereinanderliegenden endlichen Schlitzen*, wie Fig. 5 zeigt, in welcher $p = 3$ gewählt ist; in Fig. ist die Ränderzuordnung (Uferzuordnung) durch Ziffern angedeutet. Es liegt hier der Fall maximalen Zusammenhanges $2p$ des Bereiches $\tilde{\Sigma}_1$ vor.

Fig. 5.
($\lambda = 3$, $p = 3$)

Fig. 6.
(Maximalzusammenhang $2p$
$p = 3$)

Der allgemeine Fall maximalen Zusammenhanges ist der durch Fig. 6 dargestellte: p *Paare von je zwei gleich*

langen übereinanderliegenden Schlitzen, die sonst beliebig in der Ebene gelegen sind. In Fig. 6 ist $p = 3$ gewählt.

Im Falle $\lambda = 4p + 1$ finden wir den Bereich $\widetilde{\Sigma}$ begrenzt von $4p$ nach der negativen Seite ins Unendliche gehenden, von Teilpunkten freien Schlitzen, die paarweise durch Parallelverschiebungen parallel der Achse des Imaginären auf einander bezogen sind; und zwar sind bei jedem Schlitzpaare durch die Zuordnung die oberen mit den unteren Ufern ausgetauscht. Wesentlich für den Bereich $\widetilde{\Sigma}_1$ ist noch die oben allgemein angegebene Bedingung, daß die $4p$ unendlich fernen Ecken des Bereiches vermöge der Ränderzuordnung zu einem Zykel zusammengefaßt sein müssen, wodurch gewährleistet wird, daß bei Zusammensetzung dieser Ecken in der Ordnung des Zykels genau eine volle schlichte Umgebung des unendlich fernen Punktes gewonnen wird. Dieser Fall, den wir im folgenden in eine Reihe gleichberechtigter Unterfälle zerlegen werden, ist offenbar der mit der größten Parameterzahl, nämlich $6p$.

8. Explizite Aufstellung der allgemeinen ($6p$-parametrigen) Bereiche $\widetilde{\Sigma}_1$. Wir wollen jetzt eine Regel entwickeln, nach der alle im letzt erwähnten, allgemeinen Falle bestehenden verschiedenen Möglichkeiten zur Darstellung gebracht werden können. In diesem Zusammenhange ist die allgemeine Bemerkung wichtig, daß jedem den angegebenen topologischen Bedingungen genügende Bereich $\widetilde{\Sigma}_1$ auch tatsächlich durch Abbildung einer geeignet gewählten Riemannschen Fläche bei passend gewähltem Dipol gewonnen werden kann. Dies ergibt sich nämlich daraus, daß zu jedem den Bedingungen gemäß konstruierten Bereich $\widetilde{\Sigma}_1$ eine Funktion aufgestellt werden kann, die diesen Bereich in der Weise auf eine geschlossene Riemannsche Fläche vom Geschlecht p konform abbildet, daß dabei zugeordnete Randpunkte in identische Punkte der Riemannschen Fläche übergehen. Bei dieser Abbildung geht der unendlich ferne Eckenzykel in die Umgebung eines Punktes O der Riemannschen Fläche über, das unendlich ferne Richtungselement parallel der Achse des Reellen in ein bestimmtes in O zu fixierendes Richtungselement. Wird die $\widetilde{\Sigma}_1$-Ebene als $(U + iV)$-Ebene betrachtet, so kann nunmehr U als ein Strömungspotential auf F mit dem Dipol in O und einer Dipolachse von der markierten Richtung aufgefaßt werden.

Nach dieser Zwischenbetrachtung zur Frage der Aufstellung

aller $6p$-parametrigen Bereiche $\widetilde{\Sigma}_1$ zurückkehrend bemerken wir, daß es für diesen Zweck nichts ausmacht, wenn wir unter Einführung einer topologischen Hilfsabbildung des einfach zusammenhängenden Bereiches $\widetilde{\Sigma}_1$ an Stelle dieses von $4p$ paarweise auf einander bezogenen Schlitzen begrenzten Bereichs eine Kreisfläche mit $4p$ paarweise auf einander bezogenen getrennten Randstücken (Seiten) der Betrachtung zu Grunde legen. Es handelt sich dann um die Aufstellung aller möglichen Paarungen dieser Seiten unter Wechsel des Durchlaufungssinnes gemäß der Bedingung, daß die freigebliebenen Zwischenstücke der Peripherie, die wir auch auf Punkte („Teilpunkte") zusammengeschrumpft vorstellen können, zu einem einzigen Zykel verbunden werden sollen.

Läßt man irgend eins der $2p$ Seitenpaare weg, indem man sich die betreffenden Seiten mit den benachbarten Zwischenstücken der Peripherie je zu einem Zwischenstück vereinigt denkt, so bilden die $4p-2$ nunmehr vorhandenen Zwischenstücke notwendig zwei verschiedene Zykeln; eine Bemerkung, die sich vom topologischen Staupunkte mit der anderen deckt, daß nach Aufschneidung einer einfach berandeten Riemannschen Fläche Φ vom Zusammenhange $2p+1$ mittelst $2p$ getrennten Querschnitten zu einer einfach zusammenhängenden Fläche mit $4p$ Seiten und $4p$ Teilstücken des Randes (erstere von den Querschnitten, letztere von dem ursprünglichen Rande der Fläche Φ herrührend) letztere Stücke einen Zykel bilden, welcher sich in zwei Zykeln auflöst, wenn man die beiden Seiten eines Querschnittes zur Berandung der Riemannschen Fläche hinzufügt und die so entstandene neue Fläche φ vom Zusammenhang $2p$ durch die übrigen $2p-1$ Querschnitte aufgeschnitten denkt. Dieser Hilfsvorstellung entnehmen wir sogleich noch, daß es unter den letzteren Querschnitten einen solchen geben muß, der die beiden verschiedenen Begrenzungslinien der Fläche φ verbindet, weil anderenfalls als Endresultat der ganzen Aufschneidung unmöglich einfacher Zusammenhang resultieren kann. Dieser zweite Querschnitt bedeutet für uns, wenn wir zur bezogenen Kreisscheibe zurückkehren, folgendes: Man kann nach Weglassung des ersten Seitenpaares, wodurch zwei Randzykeln entstehen, unter den übrigen Seitenpaaren eines so wählen, daß dessen Weglassung die Wiedergewinnung eines einzigen Zykels von nunmehr $4p-4$ Randstücken bewirkt. Wir erkennen so die Möglichkeit, einen Abbau der ursprünglich seitenbezogenen Kreisscheibe vorzunehmen, bei welchem nach Wegnahme jedes weiteren

Seitenpaares die resultierenden Randstücke immer in einem oder in zwei Zykeln geordnet erscheinen.

Aus der vorangehenden Analyse ergibt sich nunmehr die Synthese aller seitenbezogenen Kreisscheiben aus $2p$ Seitenpaaren mit einem einzigen Randzykel. Wir haben sukzessive $2p$ Seitenpaare einzuführen, indem wir darauf achten, daß jedes weitere Seitenpaar die Zykelnzahl entweder von 2 auf 1 oder von 1 auf 2 bringt. Hat man, bei einer bestimmten Stufe des Konstruktionsprozesses angelangt, einen Zykel, so wird man das nächste Seitenpaar durchaus beliebig einordnen können und dabei immer zwei Zykeln erhalten. Hat man dagegen zwei Zykeln, so ist das nächste Seitenpaar so einzufügen, daß die beiden Seiten desselben auf die beiden Zykeln verteilt sind.

Es ist leicht auf diese Weise alle möglichen Anordnungen aufzustellen. Für $p = 1$ ergibt sich nur die eine in Fig. 7 angegebene Anordnung. Für $p = 2$ ergeben sich die in den Figuren

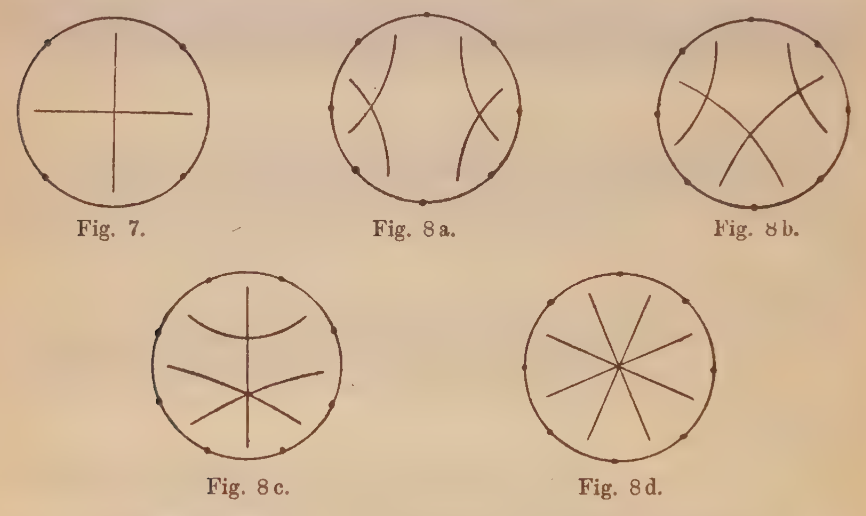


Fig. 7. Fig. 8a. Fig. 8b.

Fig. 8c. Fig. 8d.

8a—8d dargestellten 4 Anordnungen, wobei Anordnungen nur dann als wesentlich verschieden angesehen sind, wenn sie nicht durch eine stetige Verschiebung der Kreisperipherie in sich auseinander hervorgehen. In den genannten Figuren haben wir die Zwischenstücke auf Punkte zusammenschrumpfen lassen, die somit als Teilpunkte erscheinen. An Stelle des Randzykels ist infolgedessen jedesmal ein Eckenzykel getreten.

Die einzelne der Figuren 7, 8 liefert, allgemein zu reden, eine von 1 verschiedene endliche Anzahl von Schlitzbereichtypen $\tilde{\Sigma}_1$, da jetzt zu berücksichtigen ist, daß jede beliebige Seite der

bezogenen Kreisscheibe beim Übergange zum Schlitzbereich an die oberste Stelle gerückt werden kann. Es ergeben sich so aus Figur 7 ein Typ, aus Figur 8a 4 Typen, aus den Figuren 8b und 8c je 8 verschiedene Typen, aus Figur 8d 1 Typ, sodaß für $p = 2$ im ganzen 21 wesentlich verschiedene Typen $\tilde{\Sigma}_1$ der maximalen Parameterzahl $6p$ gefunden werden[1]).

9. Quadrupelförmiger Aufbau der Schlitzbereiche $\tilde{\Sigma}_1$ des allgemeinen ($6p$-parametrigen) Typus. Wir wollen zwei Seitenpaare einer in geschilderter Weise bezogenen Kreisscheibe als ein Seitenquadrupel bezeichnen, wenn diese beiden Paare sich trennen, sodaß die die Beziehung der Seiten andeutenden durch das Innere der Kreisscheibe geführten Linien sich schneiden müssen. Wir bemerken sofort, daß die zwei Seitenpaare im Falle $p = 1$ (Fig. 7) tatsächlich ein Quadrupel bilden. Ferner bemerken wir, daß im Falle $p = 2$ in jeder der vier zugehörenden Figuren 8 stets und zwar auf mehrere Weisen das System der bezogenen Seiten als die Vereinigung zweier durch passende Wahl der Seitenpaare definierten Seitenquadrupel betrachtet werden kann. Auch für beliebiges p gilt, wie wir zeigen wollen, der Satz, daß das System der $2p$ Seitenpaare einer bezogenen Kreisscheibe mit einem einzigen Randzykel stets und zwar auf solche Weise durch sukzessive Vereinigung von p Seitenquadrupeln gewonnen werden kann, daß auch in allen Zwischenstufen nur ein Randzykel da ist. Wir beweisen diesen Satz, indem wir die oben gegebene allgemeine Analyse eines solchen Systems dahin vertiefen, daß wir noch zeigen, wie der sukzessive oben dargelegte Abbau eines solchen Systems auch quadrupelweise vorgenommen werden kann. Haben wir etwa noch ein System S aus $2\varkappa$ Seitenpaaren mit einem Randzykel, so nehmen wir davon zunächst ein beliebiges Seitenpaar (s, s') heraus. Die Seiten dieses Paares vereinigen sich dadurch mit den angrenzenden Randstücken zu Zwischenstücken des übrig gebliebenen, von $2\varkappa - 1$ Seitenpaaren gebildeten Systems S'. Diese beiden Zwischenstücke gehören notwendig jedes zu einem anderen der beiden jetzt entstandenen Zykeln. Dementsprechend müssen in beiden Hälften, in die der Kreis durch das fortgelassene Seitenpaar s, s' (Fig. 9) zerlegt wird, noch Seiten von S' liegen. Die Seitenpaare von S' lassen sich nun in drei Kategorien einteilen, indem nämlich die an sie an-

1) Hilbert gibt l. c. nur einen Typus (unsere Fig. 12) an. Auch wird die Möglichkeit der Spezialtypen nicht erwähnt.

grenzenden Zwischenstücke entweder beide dem ersten Zykel oder beide dem zweiten Zykel angehören. Wir benennen die Seiten, hiernach, indem wir einen Umlaufssin im Kreise zugrundelegen, als Seiten (1, 1), (2, 2), (1, 2) bezw. (2, 1). Es ist klar, daß auf einander bezogene Seiten dieselben Benennungen tragen bis auf die Vertauschung der Benennungen (1, 2) und (2, 1). Unser Problem kommt darauf hinaus ein Seitenpaar (1, 2), (2, 1) einander entsprechender Seiten im System S' als vorhanden nachzuweisen, das mit dem weggelassenen Seitenpaar (s, s') ein Quadrupel bildet.

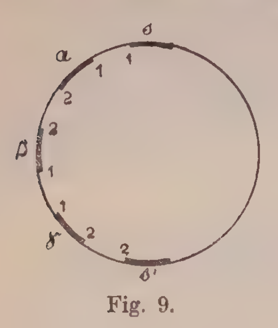


Fig. 9.

Lenken wir zu diesem Zwecke unsere Aufmerksamkeit auf die eine Kreishälfte (in Fig. 9 linke Hälfte). Aus dem Umstande, daß die Seite s in Zykel 1, die Seite s' in Zykel 2 eingeht, ergibt sich die Existenz einer Seite α der Kategorie (1, 2) auf der linken Kreishälfte. Liegt die α zugeordnete Seite β noch nicht auf der rechten Kreishälfte, so finden wir nach Einordnung der Seite β, die zur Kategorie (2, 1) gehört, bei Durchlaufung der linken Kreishälfte bei jeder Lage von β unter den Zwischenstücken zwischen s, α, β, s' eins, an dessen einem Ende Zykel 1 markiert ist, während am anderen Zykel 2 markiert ist. Es muß daher in diesem Stücke noch eine Seite γ von der Kategorie (1, 2) bzw. (2, 1) vorhanden sein, deren Bild δ wir nun wieder aufsuchen können usw. Das Verfahren muß abbrechen. Es kann nur dadurch abbrechen, daß einmal die Bildseite auf die rechte Kreishälfte fällt.

Nach dem Vorstehendem ist es möglich jedes System von Seitenpaaren einer bezogenen Kreisscheibe mit nur einem Randzykel so aufzubauen, daß man nacheinander p Seitenquadrupel wählt von der Art, daß die Zufügung jedes Seitenquadrupels die Zykelzahl unverändert auf 1 läßt. Von dieser Bemerkung kann man auch bei der praktischen Aufstellung aller möglichen Typen von $2p$ Seitenpaaren mit nur einem Randzykel Gebrauch machen. Man wird von einem Quadrupel ausgehen und $2p-1$ weitere hinzufügen, immer darauf achtend, daß ein einfacher Zykel bleibt. Um dieser Bedingung zu genügen, kann man das erste Seitenpaar eines neu hinzuzufügenden Quadrupels jedesmal willkürlich einordnen und hat dann bezüglich des zweiten Seitenpaares dieses Quadrupels auf die Erfüllung der Bedingung zu achten. Diese kann dann stets erfüllt werden.

Für unsere Schlitzbereiche ergibt sich aus dem Vorhergehenden

der Satz, daß jeder $6p$-parametrige Schlitzbereich $\tilde{\Sigma}_1$ so durch sukzessive Vereingung von p Schlitzquadrupeln gebildet werden kann, daß in jeder Stufe ein Bereich der Gattung $\tilde{\Sigma}_1$ erscheint, dessen Geschlecht natürlich gleich der Stufenzahl ist.

10. Vorkommen aller Fälle $\lambda = 3, 4, \ldots, 4p+1$ für die Zahl λ der Stromgebiete. Wir sahen oben daß die Zahl λ bei $p \geqq 1$ mindestens gleich 3, höchstens gleich $4p+1$ ist. Es erhebt sich die Frage, ob die Zahl λ alle in diesen Grenzen liegenden ganzzahligen Werte wirklich annehmen kann. Hierbei sei noch an die physikalische Bedeutung der Zahl λ als Anzahl der verschiedenen Stromverbände der durch U auf der Riemannschen Fläche F definierten Potentialströmung erinnert.

Für $\lambda = 3$ und $\lambda = 4p+1$ haben wir im Vorhergehenden nicht nur die Möglichkeit des Vorkommens festgestellt, sondern darüber hinaus alle wesentlich verschiedenen Möglichkeiten des Vorkommens ermittelt. Eine einfache Überlegung zeigt, daß die Zahl λ auch jeden Zwischenwert wirklich annehmen kann. Die

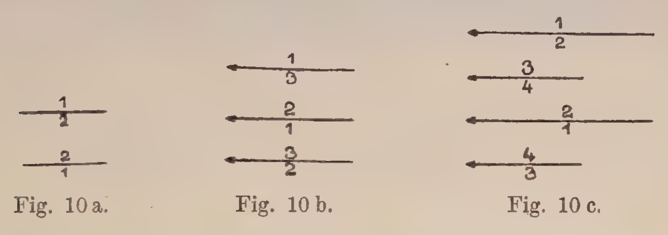


Fig. 10a. Fig. 10b. Fig. 10c.

Figuren 10 bilden eine Zusammenstellung durch welche zunächst im Falle $p = 1$ die drei Möglichkeiten $\lambda = 3, 4, 5$ veranschaulicht werden: Fig. 10a zeigt ein endliches Schlitzpaar, Fig. 10b ein aus unendlichen Schlitzen gebildetes Tripel, Fig. 10c ein aus unendlichen Schlitzen gebildetes Quadrupel. Man übersieht leicht, wie man durch Kombination (Nebeneinander- und Übereinanderstellung von Figuren der angegebenen drei Typen 10) Typen mit beliebigem p erhält und mit beliebigem λ in den Grenzen 3 und $4p+1$. Die Figurenreihe 11 veranschaulicht dieses für $p = 2$. Eine nähere Ausführung für beliebiges p erübrigt sich. Nicht beantwortet wird durch diese Figuren die Frage nach sämtlichen den Fällen $\lambda = 4, 5, \ldots, 4p$ entsprechenden Bereichtypen; (vgl. hierzu Nr. 13).

11. Verteilungsmöglichkeiten der Gesamtordnung $2p$ der Staupunkte auf die einzelnen Staupunkte. Eine Frage, die sich analog der vorhergehenden durch Exemplifikation

beantworten läßt, ist die Frage nach der Verteilung der Gesamtordnung $2p$ der Staupunkte auf die N einzelnen Staupunkte. Wir beweisen hier den Satz: es ist möglich Potentialströmungen zu realisieren, für die die Verteilung der Einheiten der Zahl $2p$ auf

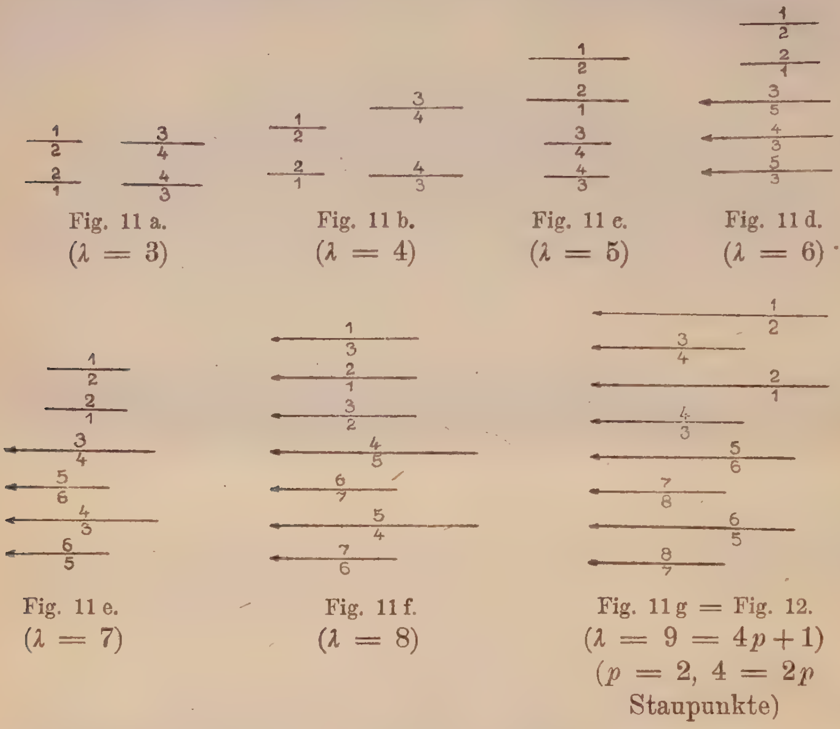


Fig. 11 a. ($\lambda = 3$) — Fig. 11 b. ($\lambda = 4$) — Fig. 11 c. ($\lambda = 5$) — Fig. 11 d. ($\lambda = 6$)

Fig. 11 e. ($\lambda = 7$) — Fig. 11 f. ($\lambda = 8$) — Fig. 11 g = Fig. 12. ($\lambda = 9 = 4p + 1$) ($p = 2$, $4 = 2p$ Staupunkte)

die Ordnungszahlen der einzelnen Staupunkte eine beliebige der Bedingung

$$(*) \qquad \tau_1 + \tau_2 + \cdots + \tau_N = 2p$$

genügende sein kann, wobei auch N eine beliebige positive ganze Zahl $\leq 2p$ sein kann.

Wir betrachten zunächst die beiden äußersten Fälle

$$N = 2p,\ \tau_1 = \tau_2 = \cdots = \tau_N = 1 \text{ und } N = 1,\ \tau = 2p.$$

Den ersten Fall realisieren wir, indem wir in der $U + iV$-Ebene p Schlitzquadrupel übereinanderstellen, wie Fig. 12 ($p = 2$) zeigt. Bilden wir den so gewonnenen Schlitzbereich auf eine Riemannsche Fläche F ab, so wird U, auf F überpflanzt, ein Strömungspotential mit $2p$ einfachen Staupunkten. Um den anderen äußersten Fall zu realisieren, setzen wir $2p + 1$ unendliche Schlitze mit übereinander stehenden Endpunkten nach einem aus den Figuren 13 und 14 ersichtlichen allgemeinen Gesetz zyklischer Anordnung

(„zyklische Schlitzsysteme ungerader Ordnung") zu einem Bereiche $\tilde{\Sigma}_1$ zusammen.

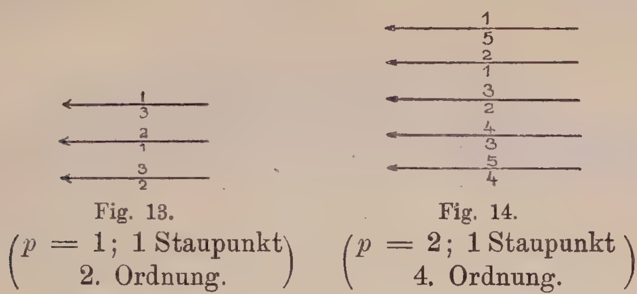


Fig. 13. $\begin{pmatrix} p = 1; \text{ 1 Staupunkt} \\ \text{2. Ordnung.} \end{pmatrix}$

Fig. 14. $\begin{pmatrix} p = 2; \text{ 1 Staupunkt} \\ \text{4. Ordnung.} \end{pmatrix}$

Zur Erledigung der Zwischenfälle der Gleichung (*) bemerken wir zunächst, daß unter den N Größen τ_α notwendig die Anzahl der ungeraden gerade sein muß, weil die Summe aller Größen τ_α gleich der geraden Zahl $2p$ ist. Die ungeraden Zahlen τ_α denken wir uns irgendwie zu zweien zusammengenommen. Es sei ν die Anzahl der geraden, $2\nu'$ die Anzahl der ungeraden Zahlen τ_α. Wir setzen dann zunächst ν „zyklische Schlitzsysteme" von der Art der Figuren 13 und 14 übereinander, deren jedes einzelne einen Staupunkt der gewünschten geraden Ordnung liefert. Weiter setzen wir $2\nu'$ „zyklische Schlitzsysteme gerader Ordnung" entsprechend den ungeraden Zahlen τ_α (je aus $\tau_\alpha + 1$ Schlitzen gebildet und folglich, da $\tau_\alpha + 1$ gerade ist, mit 2 unendlich fernen Eckenzykeln behaftet) der Reihe nach darunter, verkoppeln letztere jedoch, um einfache Zykeln im Unendlichen zu erhalten, paarweise zu einem System mit nur einem unendlich fernen Zykel, indem wir je zwei aufeinander folgende so weit gegeneinander verrücken, bis sie sich zum ersten Male überkreuzen; (vgl. Fig. 15, in welcher $p = 4$, $\tau_1 = 3$, $\tau_2 = 5$ gewählt ist).

Fig. 15. $\begin{pmatrix} p = 4; \text{ 1 Staupunkt 3. u.} \\ \text{1 Staupunkt 5. Ordnung} \end{pmatrix}$

12. **Allgemeine independente Definition der Schlitzbereiche** $\tilde{\Sigma}_1$. Um alle Schlitzbereichtypen eines bestimmten Geschlechtes p aufstellen zu können, haben wir vor allem eine independente, d. h. von der ursprünglichen Gewinnung des Schlitzbereiches aus dem Strömungspotential U unabhängige allgemeine Definition der Schlitzbereiche $\tilde{\Sigma}_1$ zu geben. Folgende Merkmale geben eine erschöpfende Charakteristik.

Ein Schlitzbereich $\tilde{\Sigma}_1$ vom Geschlecht p ist ein schlichter Bereich, der von der ganzen Ebene gebildet wird, die längs endlich vielen teils endlichen teils nach der Seite des Negativen ins Unendliche gehenden Schlitzen begrenzt wird. Die Begrenzung (— oberes und unteres Ufer eines und desselben Schlitzes gelten als verschiedene Begrenzungsteile —) ist durch Teilpunkte in endlich viele Teile zerlegt; und zwar sollen auch der unendlich ferne Punkt, so oft er als Begrenzungspunkt gesondert auftritt, sowie die endlichen Endpunkte der Begrenzungsschlitze als Teilpunkte gerechnet werden. Die Gesamtanzahl der so gewonnenen Begrenzungsteile (Begrenzungsseiten) ist gerade. Die einzelnen Seiten der Begrenzung von $\tilde{\Sigma}_1$ sind paarweise durch Parallelverschiebungen parallel der Achse des Imaginären aufeinander bezogen, wobei immer ein oberes und ein unteres Ufer zur Paarung gelangen und die identische Substritution ausgeschlossen sein soll. Die zwischen den unendlich langen Randintervallen, falls solche vorhanden sind, bestehenden Substitutionen müssen der Bedingung genügen, daß die unendlich fernen Ecken des Bereiches, deren Anzahl λ' mit der Anzahl der unendlich langen Begrenzungsschlitze von $\tilde{\Sigma}_1$ übereinstimmt, einen einzigen Zykel bilden, worin liegt, daß sich bei ihrer den Bezugsubstitutionen gemäß auszuführenden Aneinanderfügung eine schlichte volle Umgebung des unendlich fernen Punktes ergibt. Die im Endlichen liegenden Teilpunkte und nur sie geben zu weiteren Zykeln Veranlassung, deren einzelner stets einen Windungspunkt endlicher, von null verschiedener Ordnung liefert. Die Anzahl der so erhaltenen Windungspunkte ist bei richtiger Zählung ihrer Vielfachheit von selbst eine gerade Zahl $2p$, wo nun p das Geschlecht des Bereiches $\tilde{\Sigma}_1$ bedeutet, d. i. das Geschlecht der gemäß den bestehenden Seitenordnungen aus $\tilde{\Sigma}_1$ durch ideale Zusammenheftung entstehenden geschlossenen Mannigfaltigkeit. Es ergibt sich dieses durch eine im Prinzip mit der Betrachtung der No. 4 übereinstimmende Betrachtung. Die ideale durch $\tilde{\Sigma}_1$ definierte geschlossene Mannigfaltigkeit vom Geschlecht p erscheint in λ einfach zusammenhängende Gebiete zerlegt, wenn man die mit $\tilde{\Sigma}_1$ vorliegende Aufschneidung der schlichten Ebene vervollständigt, indem man alle Begrenzungsschlitze von $\tilde{\Sigma}_1$ zu beiderseits unendlichen langen geradlinigen Schlitzen verlängert.

Man bemerkt jetzt, daß bei einem Bereich $\tilde{\Sigma}_1$ vom Geschlechte

p die Anzahl der endlichen Schlitzendpunkte, vermehrt um die halbe Anzahl der übrigen endlichen Teilpunkte, die Zahl $2p$ nicht übersteigen kann, da andrenfalls die Anzahl der nach obiger Vorschrift entstehenden Windungspunkte bei Berücksichtigung deren Ordnung notwendig größer als $2p$ werden würde.

Man bemerkt auch sofort, daß die maximale Parameterzahl, nämlich $6p$, für Bereichtypen $\tilde{\Sigma}_1$ nur dann erreicht wird, wenn $\tilde{\Sigma}_1$ mit $\bar{\Sigma}$ identisch ist, d. h. wenn jeder Begrenzungsschlitz von $\tilde{\Sigma}_1$ sich als ein unendlicher, im Endlichen beginnender, sonst von Teilpunkten freier Schlitz darstellt.

Wir merken noch an, daß zufolge einer schon bei früherer Gelegenheit gemachten Bemerkung jeder den angegebenen Bedingungen gemäß konstruierte Schlitzbereich $\tilde{\Sigma}_1$ auch tatsächlich durch Abbildung einer passenden Riemannschen Fläche bei passender Wahl des Dipols gewonnen werden kann.

13. **Vollständige Aufstellung der Bereichtypen $\tilde{\Sigma}_1$ vom Geschlecht $p = 1$.** Die zuletzt gemachten Bemerkungen setzen uns in den Stand, alle Bereichtypen $\tilde{\Sigma}_1$ für beliebiges Geschlecht p durch ein Aussonderungsverfahren expizite anzugeben, was allerdings schon für $p = 2$ eine recht mühsame, doch keinerlei prinzipielle Schwierigkeiten darbietende Aufgabe ist. Für $p = 1$ ist die Aufstellung einfach. Es sind nur die durch die Figuren 16 dargestellten 6 Fälle möglich. Bei den fünfpara-

Fig. 16 a. ($p = 1$; $\lambda = 3$; $f =$ 4 Freiheitsgrade)

Fig. 16 b. ($p = 1$; $\lambda = 4$; $f = 4$)

Fig. 16 c. ($p = 1$; $\lambda = 4$; $f = 5$)

Fig. 16 d. ($p = 1$; $\lambda = 4$; $f = 5$)

Fig. 16 d. ($p = 1$; $\lambda = 4$; $f = 5$)

Fig. 16 f. ($p = 1$; $\lambda = 5$; $f = 6$)

metrigen Figuren haben wir das Vorkommnis der Vereinigung von Schlitzendpunkten mit Uferteilpunkten in einem Zykel: die beiden äußersten Endpunkte rechts liefern hier einen Windungspunkt

erster Ordnung; der übrig bleibende, stets weiter nach links gerückte dritte Schlitzendpunkt bildet zusammen mit zwei Uferteilpunkten auf den beiden anderen Schlitzen ebenfalls einen Windungspunkt erster Ordnung.

14. Independente Charakterisierung des Abbildungsproblems. Während wir bisher vom Strömungspotential U ausgingen und eine zugehörende Abbildung ermittelten, sind wir nach der independenten allgemeinen Definition des Schlitzbereiches $\tilde{\Sigma}_1$ in der Lage, auch die Abbildung selbst independent zu definieren. Das Abbildungsproblem ist das folgende: Es soll eine vorgelegte Riemannsche Fläche vom Geschlecht p in der Weise auf einen (unbekannten) Schlitzbereich $\tilde{\Sigma}_1$ eines nicht näher vorher bestimmbaren Typus abgebildet werden, daß bei der Abbildung eine gegebene Stelle O $(z = z_0)$ der Riemannschen Fläche in den unendlich fernen Punkt übergeht und gleichzeitig ein gegebenes Richtungselement in O in das unendlich ferne nach dem negativ Unendlichen gerichtete Richtungselement übergeht. Bei dieser Abbildung gehen gewisse nicht näher vorher bestimmbare Linien der Fläche F in die Begrenzungschlitze des Bereiches $\tilde{\Sigma}_1$ über und zwar jedes solche Linienstück in zwei zugeordnete Begrenzungsteile.

Die gesuchte Abbildung existiert und ist abgesehen von einer homothetischen Transformation eindeutig bestimmt; sie kann demgemäß durch Angabe der Koeffizienten A und B in der Entwicklungsform

$$\frac{A}{z - z_0} + B + ((0))$$

der Abbildungsfunktion an der Stelle O vollkommen definiert werden.

Die Differenz zweier derartigen Abbildungsfunktionen würde sich nämlich als ein Abelsches Integral erster Art mit lauter rein imaginären Periodizitätsmoduln darstellen, somit gleich einer Konstanten sein, die in Anbetracht der Normierung im Punkte O nur den Wert null haben kann.

Besonders formulieren wir noch den Satz: zwei Schlitzbereiche $\tilde{\Sigma}_1$, aufgefaßt als ideale geschlossene Riemannsche Mannigfaltigkeiten, können nur durch eine homothetische Transformation in der Weise eineindeutig aufeinander abgebildet werden, daß der unendlich ferne Punkt unter Festhaltung der von ihm ausgehenden Richtungselemente sich selbst entspricht. Bei diesem Unitätssatze

wird die Forderung, daß die Randteile der beiden Schlitzbereiche einander entsprechen sollen, nicht gestellt; ein solches Entsprechen muß sich vielmehr infolge der übrigen Abbildungsbedingungen von selbst einstellen. In letzterem Unitätssatze ist auch ausgesprochen, daß zwei Schlitzbereiche $\tilde{\Sigma}_1$ von verschiedenem Typus niemals unter Festhaltung der unendlich fernen Richtungselemente aufeinander abgebildet werden können.

Dritter Teil.

Einfache Dipolströmung und zugehörende konforme Abbildung berandeter Riemannscher Flächen.

15. Allgemeine Ermittlung der einfachen Dipolströmung und der zugehörenden Abbildung berandeter Riemannscher Flächen. Es bezeichne jetzt F eine mit q Randlinien versehene Riemannsche Fläche, die durch p nicht zerstückende Rückkehrschnitte schlichtartig gemacht werden kann. Die Zahl $2p+q$ gibt dann die Ordnung des Zusammenhanges der Fläche F an. Die Begrenzungslinien von F können wir als reguläre analytische Linien denken, ohne damit der Allgemeinheit der Untersuchung zu schaden. Wird an einer inneren Stelle O der Fläche F ein Dipol aufgesetzt, so ergibt sich ein Strömungspotential U, das analytisch durch folgende Eigenschaften bis auf eine frei bleibende additive Konstante charakterisiert ist: „U ist auf F regulär und eindeutig, abgesehen vom Punkte O, in dem U wie $r^{-1}\cos\varphi$ unstetig wird; ferner besitzt U längs der ganzen Begrenzung von F die normale Ableitung null". Die Funktion U läßt sich über die Begrenzung von F analytisch fortsetzen. Die zu U konjugierte Potentialfunktion V läßt sich ebenfalls über die Begrenzung von F hinaus analytisch fortsetzen; und zwar nimmt sie auf jeder Begrenzungslinie in jedem ihrer Zweige einen konstanten Wert an.

Zur Bestimmung der Funktion U kann man berücksichtigen, daß U über die Begrenzung von F hinweg auch auf die Rückseite der Fläche F sozusagen analytisch fortgesetzt werden kann, wobei diese Funktion an übereinanderliegenden Stellen der Ober- und Unterseite denselben Wert annimmt. Auf der so gewonnenen Doppelfläche Φ, die als eine geschlossene Fläche vom Geschlecht $2p+q-1$ anzusehen ist, ist U durch seine beiden Unstetigkeiten,

(eine auf der Oberseite und eine an der darunter liegenden Stelle der Unterseite), bis auf eine additive Konstante bestimmt. Man kann das alternierende Verfahren heranziehen, um die so auf Φ erklärte Funktion U zu konstruieren. Man kann aber auch, wenn man ein rein funktionentheoretisches Verfahren anwenden will, zunächst die Fundamentalabbildung (Uniformisierung) der Fläche F auf das Innere des Einheitkreises vornehmen und von hier aus durch Bildung Poincaréscher Θ-Quotienten oder Schottyscher Reihen[1]) eine reelle algebraische Gleichung zwischen zwei Veränderlichen aufstellen, die als eine charakteristische Gleichung der Fläche F und damit zugleich der Fläche Φ bezeichnet werden kann. Die Flächen F und Φ sind eindeutig auf die eine Hälfte bzw. das ganze durch genannte Gleichung definierte algebraische Gebilde bezogen. Die Funktion U stellt, auf das algebraische Gebilde übertragen, ein Abelsches Integral zweiter Art mit bekannten Unstetigkeiten und eindeutigem reellen Teil, d. h. mit rein imaginären Periodizitätsmoduln dar. Man kann sie daher nach bekannten funktionentheoretischen Prinzipien, ausgehend von der charakteristischen Gleichung, bilden.

Wir haben nunmehr unsere Aufmerksamkeit wieder dem Stromlinienverlauf auf der Fläche F zuzuwenden. Die Figuren 17 ver-



Fig. 17 a. Fig. 17 b.

anschaulichen den Stromlinienverlauf an einem gewöhnlichen Randpunkte. Sodann ist das Vorkommen der Randstaupunkte hervorzuheben. Die Figuren 18 zeigen das Verhalten der Strömung

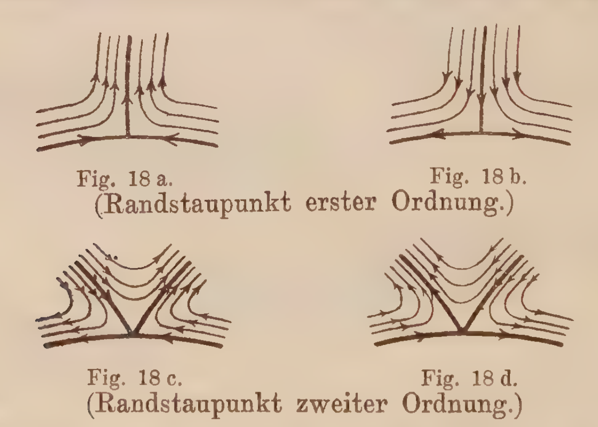


Fig. 18 a. Fig. 18 b.
(Randstaupunkt erster Ordnung.)

Fig. 18 c. Fig. 18 d.
(Randstaupunkt zweiter Ordnung.)

1) S. meine Abhandlung IV der Serie „Abhandlungen zur Theorie der konformen Abbildung“ in Acta math. Bd. 41.

in solchen Punkten. Durch Spiegelung an den Randlinien werden diese Figuren zu solchen Figuren erweitert, die inneren Staupunkten entsprechen.

Wir bemerken, daß auf jeder Begrenzungslinie der Fläche F Staupunkte vorkommen müssen, nämlich mindestens zwei verschiedene Staupunkte ungrader Ordnung, die mit solchen Randpunkten zusammenfallen, in denen die Funktion U ihr algebraisches Maximum bezw. Minimum erreicht. Im allgemeinen werden, wie aus den Entwicklungen der folgenden Nummer hervorgeht, diese Staupunkte von erster Ordnung und eindeutig bestimmt sein. Außer diesen zwei Staupunkten können aber auf derselben Begrenzungslinie noch weitere Randstaupunkte auftreten und zwar in jedem der beiden Intervalle, in welche die betrachtete Begrenzungslinie durch die beiden erwähnten Staupunkte zerlegt wird, eine grade Anzahl von Staupunkten bei Berücksichtigung ihrer Ordnung. Es hängt dieses damit zusammen, daß die Funktion U in einem Randstaupunkte ungerader Ordnung stets ein relatives Maximum oder Minimum der Randwerte annimmt. Durchläuft man also eins der genannten beiden Intervalle des Randes, so wird dabei U alle Werte von seinem minimalen bis zu seinem maximalen Randwerte annehmen. Nur beim Passieren eines Staupunktes ungerader Ordnung findet eine Unterbrechung der Monotonie der Änderung der U-Werte statt. Eine solche Änderung muß nun auf dem betrachteten Wege offenbar im ganzen eine gerade Anzahl von Malen erfolgen, sofern sie überhaupt einmal eintritt.

Nunmehr haben wir nach Analogie der entsprechenden früheren Entwicklungen das Stromverzweigungsnetz zu konstruieren, gebildet von allen Stromlinien, welche, von inneren Staupunkten oder Randstaupunkten ausgehend, das Innere der Fläche F durchziehen (singuläre Stromlinien). Wie früher zeigt man, daß das Stromverzweigungsnetz die Fläche F in eine endliche Anzahl von Stromgebieten zerlegt, in deren jedem ein Stromverband verläuft, der auf der einen Seite bei O in die Fläche F einströmt, um von der entgegengesetzten Seite bei O wieder abzuströmen. Lassen wir alle singulären Einströmungslinien weg, so gewinnen wir ein einfach zusammenhängendes Gebiet F', daß durch die Funktion $U + iV$ auf eine aufgeschlitzte Vollebene abgebildet wird. Lassen wir außerdem alle diejenigen singulären Stromlinien weg, längs denen die in F' eindeutige Funktion V keinen Peridizitätsmodul hat, so erhalten wir eine, allgemein zu reden, mehrfach zusammenhängende Fläche $\tilde{F}$, die durch die

Funktion $U + iV$ auf einen *Schlitzbereich* $\tilde{\Sigma}_2$ des gleichen Zusammenhanges abgebildet wird. Dieser Schlitzbereich spielt hier dieselbe Rolle, wie der Schlitzbereich $\tilde{\Sigma}_1$ bei der Abbildung geschlossener Flächen. Die zuletzt weggelassenen singulären Stromlinien sind, physikalisch gesprochen, diejenigen, längs denen eine Wiedervereinigung solcher Stromverbände erfolgt, die bei der ursprünglichen Einströmung benachbart waren. Der Bereich $\tilde{\Sigma}_2$ hat gegenüber dem Bereiche $\tilde{\Sigma}_1$ die Besonderheit des Vorkommens *unbezogener Begrenzungsteile*, die vermöge der im übrigen bestehenden Seitenzuordnungen, allgemein zu reden, zu *Randzykeln* als den Bildern der Randlinien von F vereinigt werden. Die Anzahl dieser Randzykeln ist gleich q, d. h. gleich der Anzahl der verschiedenen Begrenzungslinien der Fläche F.

16) *Hervorhebung einiger Haupttypen der Gattung* $\tilde{\Sigma}_2$. Der allgemeine Fall des Eingehens unbezogener Randteile in die Begrenzung des Bereiches $\tilde{\Sigma}_2$ wird der sein, welcher sich ergibt, wenn man annimmt, daß jede einzelne Randlinie der Fläche F nur zwei Staupunkte enthält, deren jeder von erster Ordnung ist, daß ferner die von diesen $2q$ Randstaupunkten ausgehenden in das Innere der Fläche führenden singulären Stromlinien keine weiteren Staupunkte treffen, sondern direkt nach O führen. Bei solcher allgemeinen Annahme muß der Bereich $\tilde{\Sigma}_2$ offenbar q *freiliegende unbezogene Begrenzungsschlitze* aufweisen.

Es findet sich so als Bereichform $\tilde{\Sigma}_2$ *größter Parameterzahl* nämlich der Parameterzahl $6p + 3q$, ein Bereich, der aus einem Bereiche $\tilde{\Sigma}_1$ größter Parameterzahl entsteht, wenn man diesen längs q getrennten endlichen Schlitzen aufschneidet. $q + 1$ ist die Ordnung des Zusammenhanges eines solchen Bereiches $\tilde{\Sigma}_2$. Dieser Bereichform entsprechend finden sich auf der Fläche F außer den erwähnten $2q$ Randstaupunkten erster Ordnung noch $2p$ getrennte innere Staupunkte erster Ordnung. Als Gesamtordnung der Staupunkte ergibt sich die Zahl

$$\sigma = 2p + q$$

wenn man, was zweckmäßig ist, die Ordnung eines Randstaupunktes bei dieser Zählung nur mit *halben Einheiten* berücksichtigt.

Bei solcher Auffassung ist $2p + q$ *ganz allgemein* die Gesamtordnung der Staupunkte unserer Strömung auf der Fläche F.

Zur Begründung werde das Differential $dW = dU + idV$ auf der Doppelfläche Φ betrachtet. Die Fläche Φ hat das Geschlecht $2p + q - 1$. Denken wir uns die Fläche Φ eineindeutig auf eine orthosymmetrische Riemannsche Fläche Φ' desselben Geschlechtes abgebildet (was dem Übergange von der Fläche Φ zu einer zu ihr gehörenden charakteristischen Gleichung (s. Nr. 15) entspricht), so wird in der Übertragung auf die z-Ebene als Ebene der Riemannschen Fläche Φ' die Größe $\frac{dW}{dz}$ eine eindeutige Funktion auf Φ', die ebenso oft null wie unendlich wird, woraus folgt, daß die Anzahl der Nullstellen vermindert um die Anzahl der Unendlichkeitsstellen des Differentials dW, d. i. die Gesamtordnung dieses Differentials, gleich der Gesamtordnung des Differentials dz ist, sofern dieses als Differential auf Φ' aufgefaßt wird. Die Gesamtordnung des letzteren ist aber gleich $w - 2m$, wenn w die Anzahl der Windungspunkte, m die Blätterzahl der Fläche Φ' bezeichnet. Nach Riemanns Analysissitus-Formel ist nun

$$w - 2m = 2(2p + q - 1) - 2,$$

da $2p + q - 1$ das Geschlecht der Fläche Φ' ist. Es findet sich daher, als Gesamtordnung des Differentials dW der Ausdruck $4p + 2q - 4$. Weil dW zwei Unendlichkeitsstellen zweiter Ordnung hat, so ergibt sich nunmehr $4p + 2q$ als Anzahl der Nullstellen des Differentials dW auf Φ, daher $2p + q$ als Gesamtordnung der Staupunkte auf F, sofern, wie oben gesagt, die Ordnung von Randstaupunkten nur mit halben Einheiten bei der Zählung der Gesamtordnung eingeführt wird.

Unter Bezugnahme auf frühere Figuren, denen wir jetzt nur noch q unbezogene Schlitze hinzuzufügen haben, ergeben sich sofort folgende Sätze:

a) Die Maximalzahl des Zusammenhanges des Bereiches $\tilde{\Sigma}_2$ ist $2p + q$.

b) Man kann durch passende Wahl einer Riemannschen Fläche F des Charakters (p, q) und passende Wahl eines Dipols auf derselben erreichen, daß $2p$ innere und $2q$ Randstaupunkte vorhanden sind, welche letztere in Paaren von je 2 getrennten Staupunkten erster Ordnung auf die q Begrenzungslinien verteilt sind, während sich für die $2p$ inneren Staupunkte eine beliebige Verteilung der Gesamtordnung im Sinne der Gleichung

$$\tau_1 + \tau_2 + \cdots + \tau_N = 2p$$

ergibt, wobei N, d. i. die Anzahl der wirklich verschiedenen inneren

Staupunkte eine beliebige positive ganze Zahl unter $2p$ (incl. $2p$) sein kann.

c) Die Zahl λ der Stromverbände kann durch passende Wahl der Bedingungen (Riemannsche Fläche, Lage und Richtung des Dipols) jeden beliebigen Wert zwischen den Grenzen 3 und $4p+q+1$ (einschl. dieser Grenzen selbst) annehmen. Eine größere Zahl von Stromverbänden ist unmöglich. Ist p gleich null, hat man es also mit einer schlichtartigen Riemannschen Fläche zu tun, so kann λ auch den Wert 2 erhalten (man denke an den Bereich, der von der längs q in ein und derselben Geraden liegenden Strecken aufgeschlitzten Ebene gebildet wird).

d) Man kann durch passende Wahl einer Riemannschen Fläche F des Charakters (p, q) und passende Wahl eines Dipols auf derselben eine beliebige Verteilung der Gesamtordnung $2p+q$ der Staupunkte erzielen, sofern diese Verteilung nur der „Bedingung" genügt, daß auf jede Randlinie eine von null verschiedene grade Anzahl von Staupunkten ungrader Ordnung entfällt. Vorgebbar ist dabei auch die Anordnung der Staupunkte mit ihren Ordnungszahlen auf den Randlinien sowie auch der Sinn der Strömung längs denselben. In der Zahl $2p+q$ als Gesamtordnung der Staupunkte sind, wie erwähnt, die Ordnungszahlen der Randstaupunkte nur mit halben Einheiten vertreten. Nach dem Sinne der vorstehenden Behauptung kann man z. B. verlangen, daß alle Staupunkte auf den Randlinien liegen.

Daß die genannte „Bedingung" notwendig ist, ergibt sich aus der Eindeutigkeit des gesuchten Potentials U. Auf jeder Randlinie ist die Gesamtordnung der auf ihr liegenden Staupunkte bei Vollzählung der einzelnen Ordnungszahlen gleich der Anzahl der auf diese Randlinie auftreffenden singulären Stromlinien. Diese singulären Stromlinien sind im Zyklus ihrer Reihenfolge abwechselnd Zuströmungs- und Abströmungslinien.

Zum Beweise der Behauptung d) wenden wir eine Methode an analog derjenigen, deren wir uns früher bei Untersuchung der entsprechenden Frage für geschlossene Riemannsche Fläche bedienten. Wir ordnen jeder Randlinie gemäß der auf ihr vorgeschriebenen Staupunktsverteilung ein besonderes zyklisches Schlitzsystem zu, bestehend aus s Schlitzen, wenn $2s$ die Anzahl der auf die Linie auftreffenden singulären Stromlinien ist. Diese Zahl s bedeutet zugleich die Gesamtordnung der Staupunkte auf der Randlinie bei Halbzählung der einzelnen Ordnungszahlen.

Wir bezeichnen als positiven Umlaufsinn der Randlinien denjenigen, bei welchem der Bereich zur Rechten liegt und unter-

scheiden darnach zwischen positiv und negativ gerichteten Randintervallen, jenachdem die Strömung in diesen Intervallen in posisivem oder negativem Sinne stattfinden soll. Den Intervallen positiver Richtung ordnen wir Längen $\alpha_1, \alpha_2, \alpha_3, \ldots$ zu, den Intervallen negativer Richtung Längen $\beta_1, \beta_2, \beta_3, \ldots$ und genügen dabei der Bedingung $\alpha_2 + \alpha_2 + \cdots = \beta_1 + \beta_2 + \cdots$. Wir numerieren nunmehr entsprechend ihrer zyklischen Anordnung die s auf die Randlinie auftreffenden singulären Stromlinien, auf denen Strömung nach dem Rande zu stattfinden soll; (in Fig. römische Ziffern). Das der Randlinie zuzuordnende Schlitzsystem entsteht dann durch ein einfaches Übertragungsverfahren, das man sich am besten an Hand der Figuren 19, 20 und 21 klarmacht. Ist s ungrade, so bilden

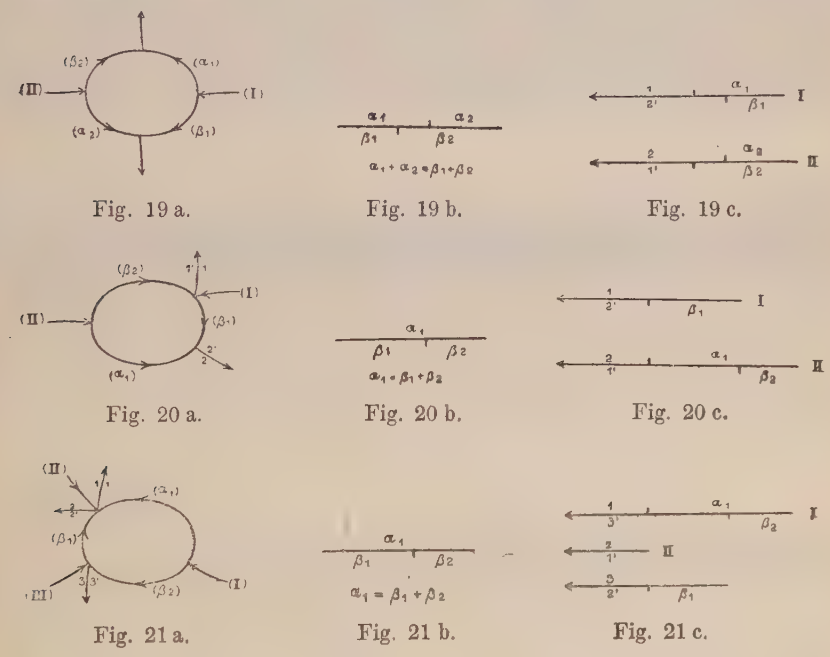


Fig. 19 a. Fig. 19 b. Fig. 19 c.
Fig. 20 a. Fig. 20 b. Fig. 20 c.
Fig. 21 a. Fig. 21 b. Fig. 21 c.

die s unendlich fernen Ecken des konstruierten Schlitzbereichs einen einzigen Zykel. Ist dagegen s gerade, so bilden sie zwei Zykeln. Man überlegt sich jedoch leicht, daß die zweizykligen Schlitzsysteme, die von Randlinien mit gradem s und inneren Staupunkten ungrader Ordnung herrühren, zusammen stets in grader Zahl vorhanden sein müssen wegen der Gesamtordnung $2p+q$ aller Staupunkte. Ist nämlich ϱ die Anzahl der Randlinien mit gradem s, so ist $q-\varrho$ die Anzahl der Randlinien mit ungradem s, folglich die Gesamtordnung der inneren Staupunkte kongruent $2p+q-(q-\varrho) \equiv \varrho \pmod 2$. Es ist also möglich die zweizykligen

Schlitzsysteme paarweise in verschränkter Lage zu je einem einzykligen System zu kombinieren (wie in Fig. 15).

Vierter Teil.

Logarithmische und polare Unstetigkeiten beliebiger Ordnung und in beliebiger Anzahl. Unstetigkeiten am Rande.

17. Zwei entgegengesetzt gleiche logarithmische Unstetigkeiten (Quellpunkt und Senkpunkt). Wir betrachten jetzt den Fall zweier getrennter auf F für das Potential U gegebener logarithmischer Unstetigkeiten $\pm \log \frac{1}{r}$, die den Punkten O_1 bzw. O_2 zugeteilt seien. Die Fläche F kann dabei als geschlossen oder berandet angenommen werden. Die Zahlen p und q sollen die alte Bedeutung haben. Physikalisch haben wir unter diesen Voraussetzungen eine Strömung, für die O_1 Einströmungsstelle, O_2 Abströmungsstelle ist. Durch Betrachtungen, die den früheren völlig analog verlaufen, finden wir nunmehr folgendes:

Die Strömung hat $2p$ bzw. $2p+q$ Staupunkte, wobei die Ordnungszahlen der Randstaupunkte nur mit halben Einheiten gezählt werden. Durch Einführung der von den Staupunkten ausgehenden singulären Stromlinien wird ein Stromverzweigungsnetz gewonnen. Dieses zerlegt die Fläche F in λ einfach zusammenhängende Stromgebiete, deren jedes einzelne in O_1 und O_2 je eine und nur eine Ecke hat. Die in O_1 beginnenden singulären Stromlinien werden nun beseitigt bis auf eine derselben, die beliebig gewählt werden kann. Auch sie wegzulassen, ist unzulässig, weil dann die Funktion $U+iV$ in dem gewonnenen einfach zusammenhängenden Bereiche F' nicht mehr eindeutig sein könnte. In der Tat würde ja einer Umkreisung der Stelle O_1 in positivem Sinne eine Vermehrung der Größe V um $2\pi i$ entsprechen. Die Fläche F' wird durch Vermittelung der Funktion $U+iV$ auf einen Parallelstreifen Σ_3 von der Breite 2π abgebildet, der parallel der Achse des Reellen gerichtet ist und in gewisser Weise durch $\lambda-1$ unendlich lange im Endlichen beginnende nach der Seite des negativ Unendlichen führende Schlitze parallel der Achse des Reellen aufgeschnitten ist. Aus F' werden jetzt noch diejenigen singulären Stromlinien entfernt, längs denen V keinen Periodizitätsmodul hat. Dadurch entsteht aus F' eine Fläche $\widetilde{F}$, die,

wenn $q > 0$ ist, im allgemeinen mehrfach zusammenhängend sein wird. Die Fläche $\tilde{F}$ wird durch $U+iV$ auf einem Schlitzbereich $\tilde{\Sigma}_3$ abgebildet, der aus $\overline{\Sigma}_3$ durch Weglassung derjenigen Begrenzungsteile entsteht, die durch die identische Substitution bezogen sind.

Bei der dargelegten Auffassung sind die obere und untere Begrenzungsgerade des Bereiches $\tilde{\Sigma}_3$ nur teilweise aufeinander bezogen, nämlich jedenfalls längs ihren der Seite des Positiven zugewandten unendlichen Enden. Man kann eine Änderung so treffen, daß die Beziehung der oberen und unteren Begrenzungsgeraden des Bereiches aufeinander eine vollständige wird. Dies wird erreicht, wenn man die oben noch beibehaltene eine singuläre in O_1 beginnende Stromlinie gleichfalls fortnimmt und sie durch irgend eine von O_1 nach O_2 führende nicht singuläre Stromlinie L ersetzt. Dieser Abänderung entspricht eine Abänderung des Bereiches $\tilde{\Sigma}_3$ zu einem Bereiche $\tilde{\Sigma}_3'$, (einen solchen Bereich zeigt Fig. 22). Man bemerkt sofort, daß zur allgemeinen Charakteri-

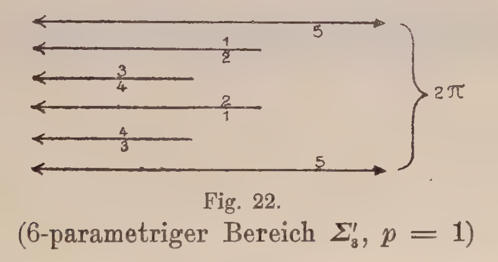


Fig. 22.
(6-parametriger Bereich Σ_3', $p = 1$)

sierung eines Bereiches $\tilde{\Sigma}_3'$ folgende Angaben ausreichen. Ein Bereich $\tilde{\Sigma}_3'$ liefert nach Fortlassung der beiden aufeinander bezogenen oberen und unteren Begrenzungsgeraden einen Bereich $\tilde{\Sigma}_1$ oder $\tilde{\Sigma}_2$, wie wir sie bei einer Dipolunstetigkeit hatten.

Es ist naheliegend die Funktion

$$W^* = e^{U+iV}$$

zu bilden. Diese Funktion wird in O_1 unendlich, in O_2 null von erster Ordnung. Sie ist in diesen Punkten nicht mehr verzweigt. Die Stromlinie L kann daher jetzt ebenfalls fortgelassen werden. Anstelle der Fläche $\tilde{F}$ bzw. $\tilde{F}'$ tritt alsdann eine Fläche $\tilde{F}^*$, die durch die Funktion W^* auf die volle W^*-Ebene abgebildet wird, nämlich auf einen Bereich $\tilde{\Sigma}_3^*$, der aus der Ebene durch Aufschlitzung derselben längs geradlinigen im Nullpunkt endigenden

oder doch nach dem Nullpunkte gerichteten Strecken entsteht, deren Ufer durch Drehungssubstitutionen mit dem Nullpunkte als Fixpunkt auf einander bezogen sind, soweit sie nicht unbezogen sind. Vgl. die Figuren 23; die Figuren 23 a—23 d veranschaulichen

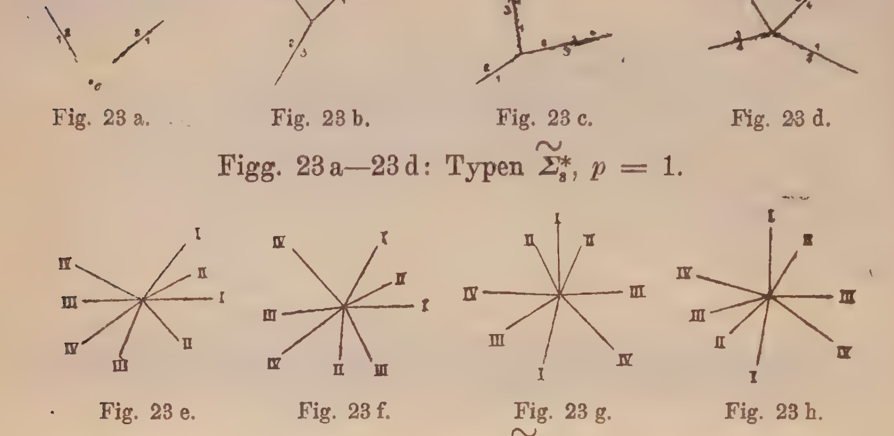


Fig. 23 a. Fig. 23 b. Fig. 23 c. Fig. 23 d.

Figg. 23 a—23 d: Typen $\widetilde{\Sigma}_3^*$, $p = 1$.

Fig. 23 e. Fig. 23 f. Fig. 23 g. Fig. 23 h.

Figg. 23 e—23 h: Allgemeine Typen $\widetilde{\Sigma}_3^*$ vom Geschlecht $p = 2$.

die verschiedenen möglichen Bereichtypen $\widetilde{\Sigma}_3^*$, die bei einer geschlossenen Fläche F vom Geschlecht 1 möglich sind, die Figuren 23 e—23 h alle möglichen Bereichtypen $\widetilde{\Sigma}_3^*$ vom Geschlecht 2 mit der maximalen Parameterzahl.

18. Endliche Anzahl von Dipolen erster Ordnung. Es seien $O_1, O_2, \ldots, O_\delta$ von einander verschiedene Dipole im Inneren der Fläche F, die geschlossen oder auch berandet sein kann. Die Zeichen p und q mögen die alte Bedeutung für die Fläche F haben. Im Punkte $O_\varkappa\,[\varkappa = 1, 2, \ldots, \delta]$ sei eine Unstetigkeit der Form

$$C_\varkappa r^{-1} \cos(\varphi - \varphi_\varkappa) \qquad (C_\varkappa > 0)\,^{1)}$$

gegeben. Das zugehörende Potential U ist auf F als ein eindeutiges Potential definiert, das bei vorkommender Berandung der Fläche F längs dieser die normale Ableitung null hat. Liegt der Spezialfall $p = 0$ vor, haben wir es also mit einer schlichtartigen Fläche F zu tun, so ist auch V ein eindeutiges Potential auf F, mithin $U + iV$ eine eindeutige Funktion auf F mit δ Unendlichkeitsstellen erster Ordnung. Letztere Funktion leistet dann eine eineindeutige konforme Abbildung der Fläche F auf eine Fläche $\widetilde{\Sigma}_4$, die sich

1) Negatives $C_\varkappa$ würde durch Ersetzung von $\varphi_\varkappa$ durch $\varphi_\varkappa + \pi$ in den entgegengesetzt gleichen positiven Wert verwandelt werden können.

als geschlossene δ-blättrige Fläche darstellt, wenn $q = 0$ ist, d. h. F selbst geschlossen ist, andernfalls, nämlich bei Zugrundelegung der Annahme $q > 0$, als eine die Ebene allenthalben δ-blättrig bedeckende Fläche, deren jede einzelne Begrenzungslinie ganz über einer endlichen Strecke parallel der Achse des Reellen verläuft. Die den Fall $q > 0$ betreffende Behauptung ergibt sich sofort aus der angegebenen Natur der Begrenzungslinien, wenn man bedenkt, daß die genannten Strecken die Ebene nicht in mehrere getrennte Gebiete zerlegen, sodaß die Bedeckung überall gleichvielblättrig ausfallen muß, nämlich δ-blättrig, weil der unendlich ferne Punkt δ-es bedeckt wird[1]).

Wir kehren jetzt zur allgemeinen Betrachtung ($p \geqq 0$) zurück. Als Anzahl σ der Staupunkte ergibt sich

$$\sigma = 2p + 2\delta - 2 \text{ bzw. } \sigma = 2p + 2\delta - 2 + q,$$

wobei Randstaupunkte bei der Zählung nur mit halben Einheiten eingehen. Das Stromverzweigungsnetz wird in der alten Weise definiert. Von jedem Punkte $O_\varkappa$ müssen singuläre Stromlinien ausgehen. Gleichfalls müssen in jedem Punkte $O_\varkappa$ singuläre Stromlinien endigen. Ginge z. B. von O_1 keine singuläre Stromlinie aus, so würden sämtliche in O_1 beginnenden Stromlinien in ihrem ganzen Verlaufe glatt neben einander her laufen. Infolgedessen würde die Funktion $U + iV$ den von diesen Stromlinien überstrichenen Teil der Fläche F auf die unaufgeschnittene ganze schlichte Ebene abbilden, was offenbar nicht zulässig ist.

Die Zahl der im Punkte $O_\varkappa$ anfangenden, ($\mu_\varkappa^+$), und endigenden, ($\mu_\varkappa^-$), singulären Stromlinien kann dabei den kleinsten Wert 1 erhalten. Aus diesen Bemerkungen ergibt sich bereits, daß die Zahl λ der Stromverbände mindestens gleich 2δ sein muß. Daß diese kleinstmögliche Zahl tatsächlich vorkommt und zwar bei beliebigem p und q, ergibt sich folgendermaßen. Man konstruiere in einer W-Ebene eine δ-blättrige Riemannsche Fläche vom Geschlecht p, deren sämtliche Windungspunkte auf der Achse des Reellen liegen, und denke sich diese Fläche evtl. noch mit q Schlitzen versehen, die ebenfalls über der Achse des Reellen gelegen sind. Wird nunmehr $W = U + iV$ gesetzt, so ist U auf der angegebenen Riemannschen Fläche ein Potential mit den verlangten Eigenschaften. Die Zahl der Stromverbände ist gleich 2δ. Die 2δ Stromgebiete werden offenbar von den 2δ Halbebenen gebildet, in die die Fläche F

1) Wegen dieses Spezialresultates vgl. F. Cecioni in Rend. del Circ. mat. di Palermo 1907.

zerfällt, wenn man sie durch alle Blätter hindurch längs der ganzen Achse des Reellen aufschneidet.

Die Zahlen $\mu_\varkappa^+$ und $\mu_\varkappa^-$ brauchen übrigens, wie wir noch hervorheben wollen, keineswegs einander gleich zu sein; dieses war wohl für $\delta = 1$ der Fall. Jetzt besteht die aus den früheren ähnlichen Gründen analoge Gleichung

$$\sum_{\varkappa=1}^{\delta} \mu_\varkappa^+ = \sum_{\varkappa=1}^{\delta} \mu_\varkappa^-,$$

d. h.: Die Gesamtanzahl aller in den Punkten $O_\varkappa$ beginnenden singulären Stromlinien ist gleich der Gesamtzahl aller in diesen Punkten endigenden singulären Stromlinien.

Kehren wir zur allgemeinen Betrachtung zurück. Wir beseitigen aus dem Stromverzweigungsnetz alle in Punkten $O_\varkappa$ beginnenden singulären Stromlinien. Hierdurch wird die Fläche F in δ einfach zusammenhängende Gebiete $G_\varkappa$ zerlegt, deren einzelnes durch die Funktion $U + iV$ auf eine aufgeschlitzte Vollebene abgebildet wird. Die Schlitze erstrecken sich dabei jeder von einem endlichen Punkte der Ebene aus nach der Seite des Negativen ins Unendliche.

Das Gebiet G_1 muß auf F an mindestens eins der übrigen Gebiete $G_\varkappa$ längs mindestens einer singulären Stromlinie anstoßen. Es sei G_2 ein solches Gebiet. Wir beseitigen dann eine solche singuläre Stromlinie und stellen so eine Verbindung von G_1 mit G_2 her. Das gewonnene Gebiet $G_1 + G_2$ ist einfach zusammenhängend. Es muß an mindestens eins der übrigen Gebiete $G_\varkappa$ längs mindestens einer singulären Stromlinie anstoßen. Es sei G_3 ein solches Gebiet. Wir erweitern $G_1 + G_2$ nunmehr zu dem Gebiete $G_1 + G_2 + G_3$, indem wir wieder eine singuläre Stromlinie beseitigen. So gelangen wir schließlich zu einem Gebiete

$$G_2 + G_2 + \cdots + G_\delta = \overline{F},$$

das ebenso wie alle $G_\varkappa$ einfach zusammenhängend ist und nun offenbar durch die Funktion $U + iV$ auf ein von δ übereinander gehefteten Vollblättern gebildetes Riemannsches Flächenstück Σ_4 mit teils auf einander bezogenen teils unbezogenen Randteilen abgebildet wird. Lassen wir schließlich noch solche singulären Stromlinien weg, längs denen die in $\overline{F}$ eindeutige Funktion $U + iV$ keinen Peridizitätsmodul besitzt, so entsteht aus $\overline{F}$ eine allgemein zu reden mehrfach zusammenhängende Fläche $\overset{\sim}{\overline{F}}$, die durch $U + iV$

auf einen δ-blättrigen Schlitzbereich $\tilde{\Sigma}_4$ mit teils bezogenen teils unbezogenen Randteilen abgebildet wird.

Es kann vorkommen, daß $\tilde{F}$ mit der geschlossenen Fläche F identisch ist. Dieser Fall liegt immer dann und nur dann vor, wenn $U+iV$ auf F eindeutig ist. In diesem Falle ist $\tilde{\Sigma}_4$ eine ebenfalls geschlossene δ-blättrige Riemannsche Fläche, nämlich das Bild von F vermöge eineindeutiger Abbildung durch die jetzt algebraische Funktion $U+iV$.

19. Endlich viele nur rein logarithmische Unstetigkeiten. Liegen endlich viele logarithmische Unstetigkeiten und nur solche vor, so sei $\mathfrak{q}^+$ die Anzahl der positiv logarithmischen Unstetigkeitsstellen (Quellpunkte), an deren $\varkappa$-ter das Potential U unstetig wird wie

$$C_\varkappa^+ \log \frac{1}{r} \qquad (C_\varkappa > 0),$$

ferner $\mathfrak{q}^-$ die Anzahl der negativ logarithmischen Unstetigkeitsstellen (Senkpunkte), an deren $\varkappa'$-ter das Potential U unstetig wird wie

$$C_{\varkappa'}^- \log \frac{1}{r} \qquad (C_{\varkappa'}^- > 0).$$

Es muß dann aus bekannten Gründen, wie sich übrigens auch weiter unten von selbst ergibt

$$(*) \qquad \sum_{\varkappa=1}^{\mathfrak{q}^+} C_\varkappa^+ = \sum_{\varkappa'=1}^{\mathfrak{q}^-} C_{\varkappa'}^- = C$$

sein. Die Anzahl σ der Staupunkte wird jetzt gleich

$$\sigma = 2p + q + \mathfrak{q}^+ + \mathfrak{q}^-$$

bei Halbzählung der Randstaupunkte. Auf jeder Randlinie muß sich eine gerade Anzahl von Randstaupunkten befinden bei Vollzählung ihrer Ordnung. Das Stromverzweigungsnetz bietet an jedem Quellpunkte und Senkpunkte mindestens eine Stromlinie dar. Nach oben angewandten Gedankengängen (Nr. 17 u. 18) gelangen wir zur Definition eines Bereichs $\tilde{F}$, der nur von solchen singulären Stromlinien begrenzt ist, längs denen $U+iV$ Periodizitätsmoduln hat; insbesondere liefert jeder Quellpunkt zur Begrenzung von $\tilde{F}$ genau eine singuläre Stromlinie. Wird diese (wie in No. 17) durch eine gewöhnliche Stromlinie ersetzt, so tritt ein Bereich $\tilde{F}'$ an Stelle des Bereichs $\tilde{F}$. Der Bereich $\tilde{F}$ bezw. $\tilde{F}'$ wird durch

die Funktion $U+iV$ allgemein zu reden auf eine Fläche $\tilde{\Sigma}_5$ bezw. $\tilde{\Sigma}'_5$ abgebildet, die aus $\mathfrak{q}^+$ in gewisser Weise übereinandergehefteten streifenförmigen Blättern gebildet wird. Der $\varkappa$-te Quellpunkt liefert einen solchen Streifen von der Breite $2\pi C_\varkappa^+$. Da jeder Stromverband in einem Quellpunkte beginnt und in einem Senkpunkte endigt, hat man sofort die schon oben erwähnte Gleichung (*).

Durch passende Wahl der bei der Bildung der Fläche $\tilde{F}$ wegzulassenden singulären Stromlinien kann unter Umständen erreicht werden, daß es nicht zu einer Übereinanderheftung, sondern zu einer teilweisen oder vollständigen Nebeneinanderheftung der streifenförmigen Blätter kommt. Bei vollständiger Nebeneinanderheftung ergibt sich so ein in gewisser Weise aufgeschlitzter parallelstreifenförmiger schlichter Bereich $\tilde{\Sigma}_5$ von der Breite $2\pi C$.

Beispiel: Als Beispiel des soeben erwähnten Vorkommnisses betrachten wir die Abbildung der schlichten z-Ebene ($p = 0$) als Fläche F durch die Funktion

$$W = U + iV = \sum_{\varkappa=1}^{\mathfrak{q}^+} C_\varkappa^+ \log \frac{1}{z-\alpha_\varkappa} - \sum_{\varkappa'=1}^{\mathfrak{q}^-} C_{\varkappa'}^- \log \frac{1}{z-\beta_{\varkappa'}}$$

in der die $\alpha_\varkappa$ und $\beta_{\varkappa'}$ reelle Größen sind und zwar

$$\alpha_1 > \alpha_2 > \cdots > \alpha_{\mathfrak{q}^+} \qquad > \beta_1 > \beta_2 > \cdots > \beta_{\mathfrak{q}^-},$$

ferner die C sämtlich positive Größen sind, die der Bedingung

$$\sum_\varkappa C_\varkappa^+ = \sum_{\varkappa'} C_{\varkappa'}^- = C$$

genügen. Beachtet man, daß $\frac{dW}{dz}$ eine rationale Funktion des Grades $\mathfrak{q}^+ + \mathfrak{q}^-$ ist, die im Unendlichen von zweiter Ordnung verschwindet, so findet man, daß die sämtlichen endlichen Nullstellen dieser Funktion, d. s. die Nullstellen des Differentials dW, zwischen den α bezw. β liegen und zwar in jedem dieser $\mathfrak{q}^+ + \mathfrak{q}^- - 2$ Intervalle eine Nullstelle erster Ordnung. Man findet daher, als Bild der oberen z-Halbebene einen geschlitzten Parallelstreifen von der

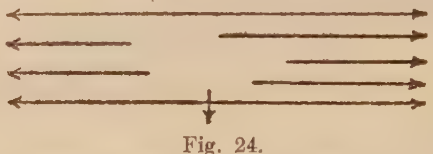


Fig. 24.

Breite πC und der Form Fig. 24, in der $\mathfrak{q}^+ = 4$, $\mathfrak{q}^- = 3$ gewählt ist. Als Bild der ganzen längs des Intervalles $\beta_1 \ldots \infty \ldots \alpha_{\mathfrak{q}^+}$ aufge-

schlitzten Ebene ergibt sich demnach ein Parallelstreifen der Breite $2\pi C$, der durch Spiegelung des erstgenannten Parallelstreifens an der unteren Begrenzungsgraden entsteht. Derjenige Teil der Achse des Reellen, der sich von α_1 über ∞ nach β_q- erstreckt, ist die einzige gewöhnliche Stromlinie, die an der betrachteten Aufschneidung der Ebene teilnimmt.

20. Pole beliebiger Ordnung. In einem Punkte O der Fläche F einen Pol der Ordnung $\mathfrak{d}$ für das Strömungspotential U geben heißt für dieses eine Unstetigkeit der Form

$$c_{\mathfrak{d}} r^{-\mathfrak{d}} \cos\{\mathfrak{d}(\varphi - \varphi^{(\mathfrak{d})})\} + c_{\mathfrak{d}-1} r^{-\mathfrak{d}+1} \cos\{(\mathfrak{d}-1)(\varphi - \varphi^{(\mathfrak{d}-1)})\} + \cdots$$
$$+ c_1 r^{-1} \cos(\varphi - \varphi^{(1)})$$

im Punkte O vorschreiben, wobei die c als positive Größen gewählt werden sollen. Die Bezeichnung der Ordnung des Poles mit $\mathfrak{d}$ soll an die Entstehungsmöglichkeit eines solchen Poles aus $\mathfrak{d}$ zusammenrückenden Dipolen passender Anordnung und Intensität erinnern. Der obige Ausdruck läßt sich auch mit komplexen Koeffizienten γ in der Form schreiben:

$$\mathfrak{R}\left[\frac{\gamma_{\mathfrak{d}}}{(z-z_0)^{\mathfrak{d}}} + \frac{\gamma_{\mathfrak{d}-1}}{(z-z_0)^{\mathfrak{d}-1}} + \cdots + \frac{\gamma_1}{z-z_0}\right].$$

Ein Dipol, wie wir ihn früher betrachtet, ist hiernach als ein Strömungspol erster Ordnung für die Funktion U zu bezeichnen.

Wir geben jetzt für die Funktion U eine endliche Anzahl δ von Polen der Ordnungen $\mathfrak{d}_1, \mathfrak{d}_2, \ldots, \mathfrak{d}_\delta$. Die Größe

$$\mathfrak{d}_1 + \mathfrak{d}_2 + \cdots + \mathfrak{d}_\delta = \mathfrak{D}$$

bezeichnet die Gesamtordnung der gegebenen Pole. Die Funktion $U + iV$ vermittelt dann eine konforme Abbildung der in gewisser Weise zu einer Fläche $\tilde{F}$ aufgeschnitten zu denkenden Fläche F auf eine von $\mathfrak{D}$ übereinandergehefteten, ihrerseits schlitzförmig begrenzten Vollblättern gebildete Riemannsche Fläche $\tilde{\Sigma}_6$ mit teils bezogenen teils unbezogenen Rändern.

Zur Begründung der vorstehenden Behauptung genügen nach den vorhergehenden Entwicklungen wenige Bemerkungen. Das Verhalten der Stromlinien in den jetzt zu betrachtenden Polen höherer Ordnung wird durch die Figur 25 veranschaulicht. Wir unterscheiden bei einem Pole der Ordnung $\mathfrak{d}$ im ganzen $\mathfrak{d}$ Einströmungsrichtungen ($U > 0$) und $\mathfrak{d}$ Abströmungsrichtungen ($U < 0$) die, miteinander abwechselnd, $2\mathfrak{d}$ gleiche Winkel

der Öffnung $\frac{\pi}{\mathfrak{d}}$ bilden. Die Anzahl σ der Staupunkte ist jetzt gleich

$$\sigma = 2p - 2 + q + \sum_{\varkappa=1}^{\delta} (\mathfrak{d}_\varkappa + 1) = 2p - 2 + q + \mathfrak{D} + \delta$$

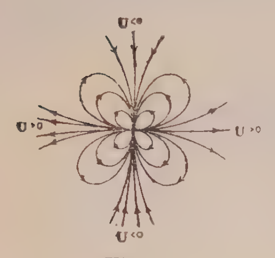


Fig. 25.
Strömungspol zweiter Ordnung.

bei Halbzählung der Randstaupunkte. Das Stromverzweigungsnetz liefert in Nachbarschaft der Pole, jeder Einströmungs- und Abströmungsrichtung desselben entsprechend, mindestens eine singuläre Stromlinie. Die Zahl λ der Stromverbände ist daher mindestens gleich $2\mathfrak{D}$. Sie kann diesen untersten Wert tatsächlich erreichen, wovon man sich wie bei einer früheren Gelegenheit (Nr. 18) durch Bildung einer jetzt $\mathfrak{D}$-blättrig zu wählenden Riemannschen Fläche mit δ Windungspunkten der Blätterzahlen $\mathfrak{d}_1, \ldots, \mathfrak{d}_\delta$ im Unendlichen überzeugt. Durch Übereinanderheftung elementarer geschlossener Windungsflächen mit je 2 Windungspunkten von den angegebenen Blätterzahlen kann man ein beliebiges Geschlecht p der Riemannschen Fläche erzielen und durch Ausführung von Schlitzen in der so gewonnenen Fläche eine beliebige Anzahl q von Randlinien.

Durch Beseitigung aller den Einströmungsrichtungen entsprechenden singulären Stromlinien ergibt sich zunächst eine Zerlegung der Fläche F in $\mathfrak{D}$ einfach zusammenhängende Bereiche, deren einzelner auf ein Vollblatt abgebildet wird. Durch Weglassung geeigneter weiterer singulärer Stromlinien gelangt man von hier aus zu einer Fläche $\widetilde{F}$, die auf einen durch passende Übereinanderheftung der Vollblätter zu gewinnenden $\mathfrak{D}$-blättrigen Bereich $\widetilde{\Sigma}_6$ abgebildet wird.

21. Gemischte Unstetigkeiten. Wir betrachten nunmehr den allgemeinen Fall, in welchem polare und logarithmische Unstetigkeiten in beliebiger Anzahl vorkommen. Insbesondere lassen wir auch zu, daß an ein und derselben Stelle polare und logarithmische Unstetigkeiten gemischt auftreten (gemischte Unstetigkeiten), d. h. Entwicklungen der Form:

$$c_{\mathfrak{d}} r^{-\mathfrak{d}} \cos \{\mathfrak{d}(\varphi - \varphi^{(\mathfrak{d})}\} + \cdots + c_1 r^{-1} \cos(\varphi - \varphi^{(1)}) \pm C \log \frac{1}{r},$$

wobei die Größen $c_1, \ldots, c_{\mathfrak{d}}$, C alle positiv angenommen werden können. Auch bei Zulassung solcher Unstetigkeiten gilt zunächst

wieder die Bemerkung, daß die Summe der Koeffizienten aller vorkommenden positiv logarithmischen Glieder gleich der Summe der Koeffizienten aller vorkommenden negativ logarithmischen Glieder sein muß. Diese Bedingung ist notwendig und hinreichend, damit ein zugehöriges eindeutiges Potential U mit der normalen Ableitung null existiere.

Für die Durchführung der Untersuchung ist vor allem der Stromlinienverlauf in der Umgebung der gemischten Unstetigkeitsstellen zu charakterisieren. Dieser wird schematisch durch dieselben Figuren erläutert, die den Stromlinienverlauf in der Umgebung rein polarer Unstetigkeiten derselben Ordnung darstellen (Fig. 25). Die Anzahl der Staupunkte wird jetzt gleich

$$\sigma = 2p - 2 + q + \sum_{\varkappa=1}^{\delta} (\mathfrak{d}_\varkappa + 1) + {}'\mathfrak{q}^+ + {}'\mathfrak{q}^-$$

$$= 2p - 2 + q + \mathfrak{D} + \delta + {}'\mathfrak{q}^+ + {}'\mathfrak{q}^-.$$

Hierbei ist mit δ die Anzahl der verschiedenen rein polaren oder gemischten Unstetigkeiten, mit $\mathfrak{D} = \mathfrak{d}_1 + \mathfrak{d}_2 + \cdots + \mathfrak{d}_\delta$ die Gesamtordnung derselben, mit ${}'\mathfrak{q}^+$ die Anzahl der positiv rein logarithmischen, mit ${}'\mathfrak{q}^-$ die Anzahl der negativ rein logarithmischen Unstetigkeiten bezeichnet.

Das von den Staupunkten aus konstruiert zu denkende Stromverzweigungsnetz wird in der alten Weise gebildet. Die singulären Einströmungslinien werden sämtlich gelöscht bis auf je eine bei den positiv rein logarithmischen Unstetigkeitsstellen. Dadurch sind verschiedene Gebiete entstanden, deren einzelnes als Bild entweder eine geschlitzte Vollebene (im ganzen $\mathfrak{D}$) oder einen Parallelstreifen (im ganzen ${}'\mathfrak{q}^+$) ergibt. Es ergibt sich somit schließlich eine Abbildung der in geeigneter Weise zu einer Fläche $\widetilde{F}$ aufgeschnitten zu denkenden Fläche F auf eine Fläche $\widetilde{\Sigma}_7$, die durch Aneinanderheftung von $\mathfrak{D}$ Vollblättern und ${}'\mathfrak{q}^+$ Parallelstreifen entsteht. Bei passender Aneinanderheftung kann es sich mitunter ereignen, daß sich im ganzen eine allenthalben $\mathfrak{D}$-blättrige Fläche $\widetilde{\Sigma}_7$ ergibt. Dieses tritt ohnehin nach der allgemein gegebenen Vorschrift immer dann ein, wenn keine positiv rein logarithmischen Unstetigkeiten vorkommen d. i., wenn ${}'\mathfrak{q}^+ = 0$ ist. Kommen keine negativ rein logarithmischen Unstetigkeiten vor, so ist $-U$ ein Potential ohne positiv rein logarithmische Unstetigkeiten. In diesem Falle würde sich also eine Abbildung auf einen allenthalben

genau $\mathfrak{D}$-blättrigen Bereich $(\tilde{\Sigma}_7)$ angeben lassen, der aus einem Bereiche $\tilde{\Sigma}_7$ durch eine 180°-Drehung entsteht.

Beispiel: Ein einfaches Beispiel für das erwähnte Vorkommnis $('\mathfrak{q}^+ = 0)$ bietet die durch die Funktion

$$W = U + iV = z + \log z$$

vermittelte konforme Abbildung der schlichten z-Ebene. Man hat in diesem Falle $p = 0$, $q = 0$

$$U = x + \log r, \quad V = y + \varphi,$$

wenn $z = r(\cos\varphi + i\sin\varphi)$ gesetzt wird. Das Potential U hat im Unendlichen eine gemischte Unstetigkeit erster Ordnung, außerdem im Nullpunkte eine negativ rein logarithmische Unstetigkeit. Demgemäß bekommen wir eine Abbildung auf die geschlitzte einblättrige Vollebene.

Es ergibt sich ein Staupunkt, nämlich der Punkt $z = -1$ aus der Gleichung $\frac{dW}{dz} = 0$. Dieser Staupunkt ist von erster Ordnung. Es gehen von ihm vier singuläre Stromlinien aus, die in Fig. 26 a stärker ausgezogen sind; die Pfeile deuten die Strom-



Fig. 26 a. Fig. 26 b.

richtung an. Die nach rechts ins Unendliche gehenden Stromlinien haben die Gleichungen $y = \pm(\pi - \varphi)$, nähern sich also asymptotisch den beiden Graden $y = \pm\pi$. Die Zahl der Stromgebiete ist gleich 2. Diese beiden Stromgebiete werden nun der allgemeinen Vorschrift entsprechend vereinigt, wenn man die letzt erwähnten zwei singulären Stromlinien, die Einströmungslinien sind, wegläßt. Die Abbildung erfolgt offenbar auf eine in der Weise wie Fig. 26 b zeigt, aufgeschlitzte Ebene.

22. Unstetigkeiten auf dem Rande. Indem wir uns jetzt noch entschließen auch Unstetigkeiten auf dem Rande zuzulassen, haben wir, um die allgemeine Form solcher Randunstetigkeiten aufzustellen, zu berücksichtigen, daß die bei der analytischen Fortsetzung eines Potentials U mit der normalen Ableitung null

am Rande über diesen hinaus sich einstellende Symmetrie den entscheidenden Einfluß auf die Form der zuzulassenden Unstetigkeiten hat. Wir denken uns eine Umgebung einer Randstelle P, an welcher eine Unstetigkeit vorgegeben werden soll, durch eine lokale konforme Abbildung auf die Fläche eines gestreckten Winkels abgebildet, der bei Einführung von Polarkoordinaten durch die Ungleichheitsbeziehungen $r' < R$, $0 \leqq \varphi' \leqq \pi$ definiert sein möge. Der Punkt $r' = 0$ entspreche dabei dem Randpunkte P. Denken wir uns die Funktion U für die Umgebung der Stelle P auf die Fläche dieses Winkelraumes überpflanzt, so muß sie nunmehr an je zwei in bezug auf die Begrenzungsgraden des Winkelraumes spiegelbildlich symmetrischen Punkten denselben Wert annehmen. Diese Eigenschaft überträgt sich auf die einzelnen Glieder der transformierten Unstetigkeit, die mithin die Form haben muß

$$c_n r'^{-n} \cos n\varphi' + c_{n-1} r'^{-n+1} \cos(n-1)\varphi' + \cdots + c_1 r'^{-1} \cos\varphi' + C \log \frac{1}{r},$$

unter $c_n, \ldots, c_1, C$ positive oder negative reelle Größen verstanden. Bei Rückübertragung auf die Stelle P ändern sich die Koeffizienten c; der Koeffizient C bleibt jedoch ungeändert, sofern angenommen wird, daß der Randpunkt P einem gewöhnlichen regulären Kurvenstücke angehöre; eine Annahme, die wir, ohne den Grad der Allgemeinheit der Untersuchung zu beschränken, als erfüllt voraussetzen können.

Das vollständige System aller im Inneren und am Rande der Fläche F vorzugebenden Entwicklungshauptteile ist jetzt nur an die Bedingung gebunden, daß die Summe der Koeffizienten der sämtlichen an inneren Punkten und Randpunkten gegebenen logarithmischen Glieder den Wert null haben muß, wobei die von Randpunkten herrührenden Koeffizienten nur mit ihrem halben Werte zu berücksichtigen sind. Ist diese Bedingung erfüllt, so gibt es stets ein, abgesehen von einer freibleibenden additiven Konstanten, völlig bestimmtes Potential U.

Für die Anzahl der Staupunkte (Randstaupunkte wieder mit halben Einheiten gezählt) ergibt sich aus der Formel der vorigen Nummer eine neue durch Hinzufügung von Gliedern, die den Randunstetigkeiten entsprechen und zur besonderen Kennzeichnung durch den Index $\mathfrak{R}$ ausgezeichnet sind, nämlich die Formel

$$\sigma = 2p - 2 + q + (\mathfrak{D} + \delta) + ('\mathfrak{q}^{+} + '\mathfrak{q}^{-}) + \tfrac{1}{2}(\mathfrak{D}_{\mathfrak{R}} + \delta_{\mathfrak{R}}) + \tfrac{1}{2}('\mathfrak{q}^{+}_{\mathfrak{R}} + '\mathfrak{q}^{-}_{\mathfrak{R}}),$$

die nun nicht einen ganzzahligen Wert σ zu liefern braucht. Der Stromlinienverlauf in den Randquellpunkten wird durch die

Figuren 27 veranschaulicht. Beseitigt man die singulären Einströmungslinien an solchen Randpunkten, so ergeben sich entsprechend den polaren oder gemischten Randunstetigkeiten nicht mehr nur Vollblättern als Teile der zu bildenden Riemannschen



Fig. 27 a. Fig. 27 b.

(Randpol erster Ordnung)

Fig. 27 c. Fig. 27 d.

(Randpol zweiter Ordnung)

Fig. 27 e. Fig. 27 f.

(Quellpunkt und Senkpunkt am Rande)

Fläche $\tilde{\Sigma}_8$, sondern auch Halbblätter (Halbebenen). Die rein logarithmischen Randquellpunkte hingegen geben wieder nur zu streifenförmigen Blättern Anlaß.

23. Der allgemeinste Bereich $\tilde{\Sigma}$ und seine physikalisch-funktionentheoretische Bedeutung. Die gefundenen Ergebnisse geben uns Veranlassung, einen allgemeinen Bereich $\tilde{\Sigma}$ zu betrachten, der durch folgende Merkmale charakterisiert ist:

Ein Bereich $\tilde{\Sigma}$ ist ein in einer W-Ebene ($W = U + iV$) ausgebreitet zu denkender endlichvielblättriger Bereich mit nur endlich vielen Windungspunkten im Innern und am Rande; er hat ferner die Eigenschaft, daß seine vollständige Begrenzung auf endlich vielen Geraden $V =$ const. enthalten ist. Zwischen je zwei benachbarten solchen Geraden $V =$ const. ist die Blätterzahl des Bereichs $\tilde{\Sigma}$ eine sich selbst gleich bleibende, während sie bei Überschreitung einer dieser Linie $V =$ const. sich ändern kann. Die Begrenzung zerfällt in endlich viele Intervalle endlicher oder unendlicher Länge, die teils unbezogene Intervalle sind teils paarweise durch Parallelverschiebungen parallel der Achse des Imaginären einander zugeordnet sind. Der vom Bereiche $\tilde{\Sigma}$ bedeckte

Teil der W-Ebene kann mit der ganzen W-Ebene identisch sein oder nur eine Halbebene oder einen Parallelstreifen in der W-Ebene bilden.

Die oben betrachteten Bereichtypen $\tilde{\Sigma}_\alpha$ bzw. $\tilde{\Sigma}'_\alpha$ $[\alpha = 1, 2, \ldots, 8]$ gehören offenbar der $\tilde{\Sigma}$-Klasse an. Man kann aber überhaupt jeden Bereich der $\tilde{\Sigma}$-Klasse in einen Bereich $\tilde{\Sigma}_\alpha$ bzw. $\tilde{\Sigma}'_\alpha$ umwandeln. Mat hat dazu seine Aufmerksamkeit den nach der Seite des Positiven ins Unendliche sich erstreckenden Teilen eines beliebig vorgelegten Bereichs $\tilde{\Sigma}$ zuzuwenden und diese Teile gemäß den bestehenden Randsubstitutionen zusammenzufügen, wobei dann soviele getrennte schlichte Gebiete entstehen, als Zykeln vorhanden sind. Bei Ausdehnung dieses Verwandlungsprozesses auf die ganze Fläche $\tilde{\Sigma}$, die man sich zu dem Zwecke zunächst durch endlich viele Geraden $V =$ const. in lauter Streifen und Halbenen zerlegt denken mag, ergibt sich eine Umordnung dieser Streifen und Halbenen in Form einer Zusammenfassung derselben entsprechend den genannten Zykeln zu schlichten Vollblättern, Halbebenen, Streifen, die ihrerseits allgemein zu reden aufgeschlitzt und nun in gewisser Weise übereinander geheftet zu denken sind.

Die Fläche $\tilde{\Sigma}$ kann nach allgemeinen Prinzipien in der Weise auf eine geschlossene oder berandete Riemannsche Fläche F abgebildet werden, daß je zwei zugeordneten Randpunkten des Bereichs $\tilde{\Sigma}$ ein und derselbe Punkt der Fläche F entspricht. Auf diese Fläche F übertragen stellt die Größe U ein Strömungspotential dar, wie wir oben betrachtet haben. Die Anwendung der erwähnten allgemeinen Prinzipien setzt die Feststellung voraus, daß die Umgebung jedes ihrer Punkte, sofern man sich die Zuordnung der Randpunkte in der Idee vorgenommen denkt, entweder auf die Fläche eines gestreckten Winkels oder auf die volle schlichte Umgebung eines gewöhnlichen Punktes abgebildet werden kann. Für die endlichen inneren oder Randpunkte des Bereichs $\tilde{\Sigma}$ ist diese Feststellung besonders einfach, da die Zusammenfügung dieser Umgebungen im Sinne der Zykelnbildung entweder zu gewöhnlichen vollen Windungsflächenstücken oder zu Winkelräumen führt, deren Öffnung ein ganzzahliges Multiplum von π ist, sodaß stets eine Transformation

$$\sqrt[\varkappa]{W - W_0} = w \text{ ($\varkappa$ ganzzahlig positiv)}$$

als lokale Hilfstransformation brauchbar ist. Bei den unendlich fernen Eckenzykeln hingegen kommt man stets mit einer Trans-

formation der Form

$$W = \frac{1}{w^{\varkappa}} + C \log w \quad (\varkappa \text{ ganzzahlig positiv, } C \text{ reell})$$

zum Ziele.

Die physikalisch-funktionentheoretische Bedeutung des allgemeinsten oben geometrisch erklärten Bereichs $\tilde{\Sigma}$ bzw. $\tilde{\Sigma}_\alpha$, $\tilde{\Sigma}'_\alpha$ läßt sich demnach so aussprechen, daß diese Bereiche vom Standpunkte des Äquivalenzbegriffs der konformen Abbildung die allgemeinsten auf einer offenen oder geschlossenen Riemannschen Fläche endlichen Zusammenhangs bei diskreter Quellenverteilung mögliche Strömung zur Anschauung bringen. Dabei ist U das Strömungspotential, wenn, wie oben, $W = U + iV$ gesetzt wird. Die Stromlinien sind die Linien $V =$ const., d. h. die Geraden parallel der Achse des Reellen. Die Stromrichtung ist die vom positiv Reellen zum negativ Reellen [1]).

1) Ich verweise zum Schluß auf eine inzwischen (Februar 1919) in der Mathematischen Zeitschrift erschienene ebenfalls den hier behandelten Gegenstand betreffende Mitteilung des Herrn R. Courant.

Translationen am Schwefel, Periklas und Kupferkies, einfache Schiebungen am Bournonit, Pyrargyrit, Kupferglanz und Silberkupferglanz.

Von

O. Mügge.

Vorgelegt in der Sitzung vom 21. Februar 1919.

Die Versuche wurden angestellt nach dem Verfahren von Johnsen[1]: Einbetten des Krystalls in ein feines Pulver (in einem stählernen Hohlcylinder) auf das mittels stählernen Stempels durch die hydraulische Presse ein Druck bis zu etwa 35000 Atm. ausgeübt wird. Als Einbettungsmaterial das in einer Flüssigkeit löslich sein muß die den Krystall und Preßcylinder nicht angreift wurde anfangs fein gepulverter Schwefel (mit Schwefelkohlenstoff als Lösungsmittel) benutzt, später kam fein gepulverter Natronsalpeter zur Anwendung da er sich gegenüber Schwefel außer durch die leichte Löslichkeit in Wasser auch durch größere Standfestigkeit auszeichnet. Da die Flächen von Schwefelkrystallen durch die Bewegung im Natronsalpeter schon ziemlich matt gerieben werden, wurde später lufttrockener Ton benutzt, der ebenfalls, auch gegenüber hohen Drucken, ausgezeichnet standhält und, auch wenn er durch starken Druck recht fest geworden ist, in Wasser leicht wieder zu Schlamm zerfällt. Als noch vorteilhafter erwies es sich, den Schwefelkrystall zunächst in etwas Feilpulver von Kautschuk oder in ein kurzes Säckchen von Gummi einzubetten, dann bleiben die Flächen ganz blank.

Die im Folgenden angegebenen Drucke haben nur die Bedeutung einer ungefähren Orientierung über die Größe des mitt-

1) Neues Jahrb. f. Mineralogie etc. B. B. **39**, 500 und 506, 1914.

leren allseitigen Druckes; um Bewegungen im Krystall zu veranlassen, darf er natürlich nicht gleichförmig sein und ist es auch nicht (nähert sich dem allerdings, wenn bei der Einbettung zu viel Gummi benutzt wird); auch ändert sich der Druck nach Größe und Richtung wenn Bewegungen im Krystall eintreten.

1. Tanslationen am Schwefel.

Hier wurden Versuche angestellt weil am rhombischen Schwefel Zwillingsbildungen nach 4 Gesetzen bekannt sind, wenn auch ihre secundäre Entstehung in keinem Falle wahrscheinlich war. Nach Pressungen bis zu 10 000 Atm. entstanden Streifen und Absonderungsflächen nach (111); stärkere (bis rd. 30 000 Atm.) bewirkten zwar weitgehende Verquetschungen der Krystalle, aber keine einfachen Schiebungen. Dabei trat Zertrümmerung nur an besonders exponierten Stellen ein, im Allgemeinen blieben die Krystalle durchaus fest. Dies Verhalten des ja als besonders spröd geltenden Schwefels ist um so bemerkenswerter als sein zum Einbetten verwendetes Pulver nach Johnsen's und meinen Erfahrungen nicht stetig, sondern explosionsartig aus den Lücken zwischen Stempel und Cylinder bei etwas höheren Drucken (15 000 bis 20 000 Atm.) zu entweichen pflegt.

Die Translationsstreifen konnten, soweit sie bündelweise auftraten über sämtliche Flächen der Krystalle (111) (113) (001) (011) verfolgt werden; da ihr Reflex nur wenig und auf derselben Fläche und für parallele Lamellen um wechselnde Beträge von dem der Hauptfläche abweicht, auch nach entgegensetzten Seiten, und an den Krystallen sowohl wie an den z. T. sehr dünnen Absonderungsblättchen nach (111) niemals Änderungen der optischen Orientierung zu erkennen waren, kann es sich nur um Translationen ∥ (111) handeln. Eine bevorzugte Translationsrichtung ist nicht mit Sicherheit anzugeben, vielleicht ist es [110], da auf (001) die Lamellen am wenigsten ausgeprägt zu sein pflegen. Längs den Translationsflächen konnte vielfach Absonderung und auch „Verwerfung" von Krystallteilen zu einander beobachtet werden, auch „Schleppung" längs der Spuren der Lamellen auf der Austrittsfläche; feine Strichelchen die manchmal trübe gewordene Krystallteile scharenweise durchziehen, rühren vermutlich daher, daß durch das Abreißen von Lamellen an ihrer Durchkreuzung mit den Absonderungsflächen linienförmige Hohlräume entstanden sind; sie erinnern sehr an ähnliche scharenweis auftretende Strichelchen in manchen Feldspaten, Augiten, Hornblenden u. a., die in Gesteinen namentlich da sich einstellen, wo sie starke Quetschungen erlitten haben.

2. Translationen am Periklas.

Dies Mineral ähnelt in seinen Cohäsionsverhältnissen durchaus dem Steinsalz: Spaltbarkeit nach (001), Translationsfähigkeit nach $T = (110)$ längs $t = [1\bar{1}0]$. Zu den Versuchen wurden nur Spaltstücke künstlicher Krystalle benutzt; erst bei Drucken oberhalb 15 000 Atm. entstanden sehr feine und scharf gradlinige Streifen; auf (001) nur parallel den Würfelkanten, auf Krystallflächen (111) nur senkrecht zu den Kanten zum Würfel. Es ist also die Translationsrichtung hier durchaus beschränkt auf $[1\bar{1}0]$. Ebenso wie beim Steinsalz kann man auf (001) Schlagfiguren von derselben Form und Orientierung und mit derselben charakteristischen Streifung ihrer Seitenflächen erzielen; sie entstehen auch, (aber fast immer nur $^1/_4$ der ganzen Figur, und zwar überall dasselbe Viertel auf derselben Fläche) an solchen Stellen der Würfelflächen die vor der Pressung schon etwas verletzt waren, indem offenbar der Salpeter hier in die kleinen Vertiefungen eindrang und nun die vorliegenden Krystallteile unter großer Reibung mit sich schleppte. Ebenso wie beim Steinsalz entstehen auch Trennungsflächen || (110) beim Pressen zwischen zwei vertikalen Kanten eines Spaltwürfels. Die Ähnlichkeit mit dem Steinsalz ist also eine vollkommene, indessen ist die Plastizität eine viel geringere; mehrfach zeigten Spaltstücke nach Drucken bis zu 20 000 Atm. keine Spur von Translationen.

3. Translationen am Kupferkies.

Obwohl am Kupferkies Zwillinge nach mehreren Gesetzen und gelegentlich an spätigen Massen auch lamellarer Zwillingsbau beobachtet ist, waren durch Druck bis zu etwa 25 000 Atm. nur Translationen hervorzurufen. Lamellen entstanden sehr zahlreich schon bei 10 000 Atm.; sie verlaufen stets || (111) und da die einzelnen Lamellen oder Schaaren von solchen über sämtliche Flächen von (111), (201), (101), (203) und (001) zu verfolgen sind, läßt sich eine bevorzugte Translationsrichtung nicht angeben. Bei stärkeren Drucken treten starke Deformationen ein, Kupferkies scheint plastischer als Schwefel, zerbröckelt aber doch viel leichter als dieser. Bestimmte Grenzflächen sind an den verschobenen Krystallteilen nicht zu erkennen und damit steht im Einklang, daß das feine Pulver gepreßter Krystalle, das bei intensiver Beleuchtung noch eben durchscheinend wird, zwischen gekreuzten Nicols optisch durchaus einheitlich erscheint.

4. Einfache Schiebungen am Bournonit.

Nicht nach (110) verzwillingte Krystalle von Bournonit kommen bekanntlich kaum vor und lamellare Wiederholung der Zwillingsbildung ist vielfach beobachtet; es war daher zu erwarten, daß Pressungen einfache Schiebungen nach (110) zur Folge haben würden und zwar bei dem pseudotetragonalen Habitus mit der zu $\varkappa_1 = (110)$ fast senkrechten Grundzone $\sigma_2 = [110]$. Die wenigen an dem kostbaren Material möglichen Versuche haben dies bestätigt. Die in den Krystallen von Horhausen erzeugten Lamellen lassen wegen ihrer Feinheit zwar Messungen nicht zu, indessen weichen ihre Grenzflächen alle in dem geforderten Sinne und in der zu erwartenden Größenordnung von ihren Hauptflächen ab, auch sind die Spuren der Lamellen auf (001) wie es das obige Verschiebungspaar verlangt, nicht zu verfolgen. Die Lamellen entstehen schon von ca. 12000 Atm. Druck an, indessen blieben auch Drucke von 20000 Atm. einige Male ohne Wirkung.

5. Einfache Schiebungen am Pyrargyrit.

Bei diesem Mineral sind Zwillingsverwachsungen nach mindestens 5 Gesetzen bekannt, von diesen erscheinen solche nach $(10\bar{1}4)$ auch lamellar und unter Umständen die eine Entstehung durch Druck wahrscheinlich machen[1]. Die Begrenzung dieser Lamellen entspricht nämlich meist derjenigen welche zu erwarten wäre, wenn

$$\varkappa_1 = (10\bar{1}4) \text{ und } \sigma_2 = [0001]$$

ist, indessen war dies, vermutlich nur in Folge späteren Weiterwachsens der Lamellen, nicht durchweg der Fall, sodaß Entscheidung durch den Versuch erwünscht schien. Dieser bestätigte das Gleitvermögen und das vermutete Verschiebungsschema.

Lamellen nach $(10\bar{1}4)$ entstehen oberhalb 16000 Atm. Druck, und zwar vielfach so zahlreich und fein, daß es schwer hält ihre Begrenzung zu ermitteln, zumal gewöhnlich Durchkreuzung der Lamellen nach mehreren Flächen $(10\bar{1}4)$, also unter sehr flachen Winkeln stattfindet, was den Eindruck erwecken kann, als ob die Fläche nur unregelmäßig gekräuselt sei; nach stärkerem Druck wurden öfter nur sehr starke (? scheinbar) unregelmäßige Deformationen beobachtet. Durch die Ebene der Schiebung $(11\bar{2}0)$ gesehen ließ sich feststellen, daß die optische Orientierung der Lamellen von der des Hauptkrystalles abweicht, wenn es auch wegen

---

1) O. M. Neues Jahrb. f. Min. 1897 II, 81.

ihrer Feinheit nicht gelang, ihre eigene optische Orientierung zu ermitteln.

Charakteristisch für die Umlagerung ist, daß sämtliche Flächen der Säulenzone ihre Indices behalten, ebenso sämtliche Flächen aus der Zone der Schiebungsrichtung, also einer Polkante von $(01\bar{1}2)$. Dabei muß man aber davon absehen, daß am Pyrargyrit wegen seiner Hemimorphie Fläche und Gegenfläche nicht gleichwertig sind und also die Hemitropie nach der Normalen der Gleitfläche nicht identisch ist mit der nach der Polkante von $(01\bar{1}2)$. Welche der beiden Zwillingsbildungen durch die einfache Schiebung bewirkt wird, konnte wegen der Schwierigkeit am Pyrargyrit Fläche und Gegenfläche zu unterscheiden, einstweilen nicht festgestellt werden.

Der Umstand, daß die einfachen Schiebungen nach allen 3 Gleitflächen $(10\bar{1}4)$ mit derselben Grundzone [0001] erfolgen, bewirkt, daß ein Krystall der zweimal die Schiebungen erst nach $\{10\bar{1}4\}$, dann nach $\{\bar{1}104\}$ und dann nach $(0\bar{1}14)$, dann wieder nach $(10\bar{1}4)$ u. s. f. erfahren hat seine ursprüngliche Form wieder erhält.

An dem dem dunkeln Rotgülden sehr ähnlichen lichten, das im Besonderen auch häufig Lamellen nach demselben Gesetz wie oben enthält, entstanden durch Druck bis zu 20000 Atm. keine Lamellen, indessen konnten nur wenige Versuche angestellt werden.

6. Einfache Schiebungen am Kupferglanz.

Hier kennt man Zwillingsbildung nach nicht weniger als 6 Gesetzen, von denen jene nach (110) und (130) direkt als mimetisch-hexagonal zu bezeichnen sind, während jene nach (011) und (112) diesem mimetischen Charakter in sofern Rechnung tragen als (011) und (112) Teile derselben pseudohexagonalen Pyramide sind. Zugleich verstärkt die Zwillingsbildung nach diesen letzteren Flächen den neben dem pseudohexagonalen Charakter vorhandenen pseudotetragonalen, indem durch sie die Fläche (001) fast mit (010) bzw. (110) zur Deckung gebracht wird.

Der mimetische Charakter der genannten 4 Zwillingsgesetze ließ erwarten, daß einfache Schiebungen, wenn überhaupt, am ehesten nach ihnen stattfinden würden, und zwar der Art, daß zu jeder Gleitfläche die zugehörige Grundzone nahezu senkrecht steht, der Betrag der Schiebung also sehr klein wird. Das ist aber nicht der Fall, vielmehr gehen Schiebungen vor sich, welche zu der in der Natur recht seltenen Zwillingsverwachsung nach (201) führen, außerdem nach einem neuen Gesetz, nämlich $\varkappa_1 = (131)$. Letztere stehen zu jenen nach $\varkappa_1 = (201)$ ebenfalls in mimetischer Bezie-

hung, indem sie beide auch Teilflächen einer pseudohexagonalen Pyramide (aber anderer Ordnung wie (011) und (112)) sind; auch die zugehörigen Grundzonen entsprechen diesem Verhältnis.

Druckversuche ergaben, daß Lamellen nach (201) schon bei ganz geringen Drucken entstehen, ja, sie lassen sich schon durch freihändigen geeigneten Druck hervorrufen; die Lamellen nach (131) entstehen vielleicht etwas schwieriger, sie wurden niemals ohne jene nach (201) erhalten.

1) Für $\varkappa_1 = (201)$ ist $\sigma_2 = [100]$,

es liegt also der bisher nicht beobachtete Fall vor, daß derselben Grundzone ein Gleitflächen**paar** zugeordnet ist[1]). Die stärkste Kippung (für (001) = zweite Kreisschnittsebene $\varkappa_2$) beträgt $33^0 24'$; das Schiebungsschema konnte durch zahlreiche Messungen der verschobenen Fläche (001) und auch anderer Flächen festgelegt werden. Der Verlauf der Zwillingsstreifen auf (110), wo sie unter nahe 30^0 zur Achse Z geneigt sind, gibt diesen Flächen eine ähnliche pseudohexagonale Streifung, wie ihn die Basis durch die oscillatorische Combinationsstreifung nach (010) in Verbindung mit dem Verlauf der Lamellen nach (110) erhält, und da beide Flächen, (110) und (001), durch die Zwillingsbildung nach (112) fast zur Deckung gelangen, können mit Zwillingsstreifung nach (201) bedeckte Säulenflächen mit nach (112) in Zwillingsstellung befindlichen Basisflächen verwechselt werden. (Ähnliches gilt von den unten besprochenen Zwillingsverwachsungen nach (131)).

Eine eigentümliche Deformation würde ein Krystall erfahren, der die beiden nach dem obigen Gesetz möglichen einfachen Schiebungen mehrfach abwechselnd nach einander erfährt. Durch das Schiebungspaar $\varkappa_1 = (201)$, $\sigma_2 = [100]$ werden die Indices einer Fläche $\{h_1 h_2 h_3\}$ verändert in $\{h_1' h_2' h_3'\}$, und es ist:

(I a) $$h_1' : h_2' : h_3' = -h_1 : h_2 : (h_3 - h_1).$$

Durch die einfache Schiebung mit den Elementen $\varkappa_1 = (20\overline{1})$, $\sigma_2 = [100]$ wird $\{h_1 h_2 h_3\}$ übergeführt in $\{h_1'' h_2'' h_3''\}$ mit

(II a) $$h_1'' : h_2'' : h_3'' = -h_1 : h_2 : (h_1 + h_3).$$

Erfährt die Fläche $\{h_1 h_2 h_3\}$ erst die Schiebung (I a), dann (II a), so erhält sie die Indices

(I a. II a) $$h_1 : h_2 : (-2h_1 + h_3).$$

1) Die Zuordnung eines Gleitflächen**tripels** zur selben Grundzone ist nach Mitteilung von Johnsen, (Centralbl. f. Min. etc. 1918, 266) von Grühn u. Veit am Magnetit, Eisenglanz u. Korund beobachtet.

Durch die n-malige Wiederholung der Schiebungen in der genannten Reihenfolge gehen die Indices von $\{h_1 h_2 h_3\}$ über in:

$$(-1)^n . h_1 : h_2 : (-n . h_1 + h_3).$$

Während also die auf die Achsen X und Y bezüglichen Indices der Fläche endlich bleiben, und für die Achse X zwischen dem ursprünglichen und dem entgegengesetzt gleichen Werte abwechseln, wird jener für die Achse Z immer größer. Jede beliebige Krystallfläche nähert sich also in ihrer Lage immer mehr der Basis (001) je öfter die Deformation wiederholt wird, d. h. der Krystall wird durch die Wiederholung beider Schiebungen immer mehr zu einer dünnen Tafel nach (001) quasi ausgewalzt ohne seine Homogenität theoretisch irgendwie einzubüßen; das Abwechseln der beiden Schiebungen nach $\{201\}$ und $\{20\bar{1}\}$ wirkt also, was die Form des Krystalls angeht, wie eine Translation nach $T = (001)$ mit $t = [100]$. Nach der 100sten Deformation würde ein Krystall der ursprünglich von je zwei parallelen Flächen (010), (001) und (201) (letztere in gleicher Größe) begrenzt war zu einem Blättchen von fast 58facher Länge parallel der Achse X und nur 1/58 der ursprünglichen Dicke ⊥ (001) (bei gleicher Breite $\| Y$).

2) Für die zweite einfache Schiebung ist

$$\varkappa_1 = \{131\}, \ \sigma_2 = [110].$$

Da diese beiden Elemente zu einander und zur Basis fast ebenso liegen wie die Elemente $\varkappa_1 = \{201\}$ und $\sigma_2 = [100]$ (indem sie die hexagonal-mimetischen Äquivalente derselben sind), ist die Kippung (für die hexagonalen Analoga der Flächen) fast dieselbe, sie konnte aber nur hinsichtlich des Sinnes der Kippung und der Größenordnung festgestellt werden, weil das gleichzeitige Vorhandensein der Lamellen nach beiden Flächen (201) und die Durchkreuzung der 4 Lamellensysteme nach (131) eine sichere Winkelmessung unmöglich machte.

Die wiederholte Vornahme zweier dieser Deformationen, etwa nach $\varkappa_1 = \{131\}$ mit $\sigma_2 = [110]$ und nach $\varkappa_1 = \{13\bar{1}\}$ mit $\sigma_2 = [110]$ (abwechselnd) würde ein ganz ähnliches Ergebnis haben wie die oben betrachteten abwechselnd nach $\{201\}$ und $\{20\bar{1}\}$. Durch die einfache Schiebung nach $\varkappa_1 = \{131\}$, $\sigma_2 = [110]$ gehen nämlich die Indices von $\{h_1 h_2 h_3\}$ über in

(Ib) $$(-h_1 + h_2) : (3 h_1 + h_2) : (h_1 + h_2 - 2 h_3),$$

durch die Schiebung nach $\varkappa_1 = \{1\bar{3}1\} \sigma_2 = [110]$ dagegen in

(IIb) $$(h_1 - h_2) : (-3 h_1 - h_2) : (h_1 + h_2 + 2 h_3).$$

Eine Fläche welche erst der Deformation (I b), dann (II b) unterliegt, erhält daher die Indices

(I b. II b) $h_1 : h_2 : (-(h_1 + h_2) + h_3)$

ebenso nach nochmaliger Deformation (I b):

(I b. II b. I b) $(-h_1 + h_2) : (3h_1 + h_2) : (3(h_1 + h_2) - 2h_3)$.

Bezeichnet man allgemein die n-malige Wiederholung der beiden Schiebungen (I b) und (II b) mit (I b. II b)^n, die dann folgende mit (I b. II b)^n. I b, so werden die Indices der Fläche:

(I b. II b)^n $h_1 : h_2 : (-2n(h_1 + h_2) + h_3)$

(I b. II b)^n. I b $(-h_1 + h_2) : (3h_1 + h_2) : ((2n+1)(h_1 + h_2) - 2h_3)$.

Während also nach der $2n$-ten Schiebung die Indices für die Achsen X und Y wieder dieselben sind wie zu Anfang, wird der Index für Z immer größer negativ; nach jeder $(2n + 1)$ Schiebung haben zwar die Indices X und Y einen anderen, aber immer denselben endlichen Wert, während der für Z immer größer positiv wird.

Auch hier nähert sich also jede beliebige Ausgangsfläche durch die Aufeinanderfolge vieler Schiebungen abwechselnd nach $\{131\}$ und $\{13\bar{1}\}$ der Basis (001), d. h. der Krystall erhält immer mehr die Form einer dünnen Tafel nach (001), welches auch seine ursprüngliche Form sein mag, seine Gestalt wird verändert als unterläge er einer immer weiter und stets in demselben Sinne fortgesetzten Translation längs (001) in der Richtung [110].

Das für den Kupferglanz gesagte gilt wahrscheinlich auch für den Silberkupferglanz. Es wurden allerdings nur die Deformationen nach $\varkappa_1 = (201)$, $\sigma_2 = [100]$ festgestellt, indessen konnte an den sehr kleinen Krystallen des kostbaren Materials nur ein Versuch angestellt werden.

Die Gitter von Rotgülden, Bournonit u. Kupferglanz.

Am Pyrargyrit ist die Grundzone [0001] dieselbe wie beim Millerit, der Schiebungen nach $\varkappa_1 = (01\bar{1}2)$ eingeht. Gibt man daher der Gleitfläche des Pyrargyrit dieselben Indices wie der des Millerit (indem man statt des jetzigen das nächst stumpfere „Rhomboeder" zum Grundrhomboeder macht), so gilt hinsichtlich der Gitter das

von Johnsen[1]) Abgeleitete, mit der Einschränkung, daß bei der unzweifelhaft trigonalen Symmetrie des P. hexagonale Gitter nicht in Frage kommen. Demnach gibt es unendlich viele Gitter die mit den Schiebungselementen des P. verträglich sind, nämlich alle mit den (Millerschen) Indices $(\pi \| \pi \pm 1 \| \pi \pm 1)$.

Für die Schiebungen rhombischer Krystalle wie Bournonit mit den Elementen $\varkappa_1 = (110)$, $\sigma_2 = [110]$ ist schon früher[2]) gezeigt, daß damit nur solche Gitter verträglich sind, deren Maschen in (001) den Spuren der beiden Kreisschnittsebenen oder deren Diagonalen parallel laufen. Damit sind in Übereinstimmung die Ergebnisse bei Anwendung der von Johnsen[3]) abgeleiteten allgemeinen Bedingungsgleichungen. Zulässig sind danach Gitter nach rechtwinkligen Parallelepipeden und centrierten derartigen Parallelepipeden deren Kantenlängen in (001) sich wie die Achseneinheiten für die Achsen X und Y des B. verhalten. Ferner auch Gitter nach rhombischen Säulen der Stellung $(hk0)$ und centrierter derartiger Säulen, wenn der Winkel der Säulen gleich dem Winkel der Säule (110) des B. (gleich dem Winkel der beiden Kreisschnittsebenen) ist. Ferner sind auch zulässig Gitter nach centrierten Säulen der Stellung $(h0l)$ wenn die Länge der Säulenkante $(\| Y)$ sich zu ihrem Durchmesser $\| X$ wie die Achseneinheiten für Y und X des B. verhalten und endlich auch centrierte Säulen der Stellung $(0kl)$ wenn das Verhältnis zwischen der Länge der Säulenkante und ihrem Durchmesser $\| Y$ des B. gleich den Einheiten für diese Achsen sind.

Ganz ausgeschlossen sind dagegen alle Gitter nach nicht centrierten rhombischen Säulen der Stellung $(h0l)$ und $(0kl)$. In allen genannten Gittern bleibt die Dimension $\| Z$ des B. unbestimmt.

Beim Kupferglanz ist die Auswahl unter den rhombischen Gittern sehr viel stärkeren Beschränkungen unterworfen (immer unter der Voraussetzung daß die beobachteten Schiebungen reine Gitterschiebungen sind). Schiebungen nach $\varkappa_1 = (201)$, $\sigma_2 = [100]$ sind zwar in allen Arten rhombischer Gitter und bei allen Stellungen der rhombischen Säulen zu den Elementen der einfachen Schiebung möglich; sie unterliegen nur der Beschränkung, daß im Gitter nach rechtwinkligen Parallelopipeden die Länge ihrer Kanten $\| Z$ des K. ein ganzes Vielfaches der Achseneinheit auf Z des K.

1) A. Johnsen, Centralbl. f. Min. usw. 1916, 128.

2) O. M. Neues Jahrb. f. Min. B. B. **14**, 311, 1901.

3) l. c. p. 121 ff.

sein muß, wenn die Länge der Kante $\|X$ gleich der Einheit dieser Achse gesetzt wird; daß ferner bei Säulen der Stellung $(0kl)$ dasselbe für das Verhältnis der Längen ihrer Durchmesser $\|Z$ zur Länge ihrer Kante $\|X$ gilt. Bei allen diesen Gittern bleibt ihre Dimension $\|Y$ unbestimmt.

Von allen diesen Gitterarten vermag aber nur eine einzige gleichzeitig die Schiebungen nach $\varkappa_1 = (131)$, $\sigma_2 = [110]$ einzugehen, nämlich nur jene nach Säulen der Stellung $(hk0)$ und auch nur dann, wenn ihr Netz in (001) entweder den Spuren der Säule (110) oder der Säule (130) des K. parallel geht und wenn gleichzeitig die Dimensionen dieser Säulen $\|$ den Achsen X und Z des K. sich verhalten wie die Einheit der Achse X zu einem graden Vielfachen der Einheit der Achse Z.

Diese für Kupferglanz möglichen Gitter sind nach ihrer Art und nach ihren Dimensionen durchaus im Einklang mit dem mimetisch-hexagonalen Habitus der Krystalle, da sie für seine Hauptflächen hohe Netzdichte wahrscheinlich machen. Außerdem aber sind sie zugleich fähig pseudohexagonale Schiebungen mit den Elementen $\varkappa_1 = (110)$, $\sigma_2 = [310]$ und $\varkappa_1 = (130)$, $\sigma_2 = [\bar{1}10]$ einzugehen wie sie bei zahlreichen rhombischen pseudohexagonalen Krystallen beobachtet sind[1]).

1) Einfache Schiebungen die zur Zwillingsbildung nach den beiden anderen Gesetzen des Kupferglanzes führen würden, nämlich nach $\varkappa_1 = (011)$ und $\varkappa_1 = (221)$ und zwar (wegen des mimetischen Charakters auch dieser Zwillinge) mit den Grundzonen $\sigma_2 = [011]$, bzw. $\sigma_2 = [312]$ (die auf den zugehörigen Gleitflächen nahezu senkrecht stehen), würden dagegen bei den für $\varkappa_1 = (201)$ und $\varkappa_1 = (131)$ abgeleiteten Gittern nicht möglich sein.

Erdbebenwellen VII.

Aus dem Nachlaß von K. Zoeppritz.

a) Nachtrag zu den Laufzeitkurven.

b) Über Reflexion und Durchgang seismischer Wellen durch Unstetigkeitsflächen.

Vorgelegt in der Sitzung vom 27. Juni 1914 von E. Wiechert.

Vorbemerkungen von E. Wiechert.

In dem wissenschaftlichen Nachlaß von K. Zoeppritz (gestorben am 20. Juli 1908) fanden sich die beiden im folgenden abgedruckten Arbeiten vor, die mir von der Witwe des Verstorbenen freundlichst für die Veröffentlichung zur Verfügung gestellt wurden. Beide Arbeiten, welche sich unmittelbar an die Veröffentlichung „Erdbebenwellen I" anschließen, enthalten vielerlei Gedanken, die von erheblicher Bedeutung für die Wissenschaft sind, sodaß die Veröffentlichung geboten scheint. Die Arbeiten sind zwar entstanden, während ich mit dem jungen Gelehrten, meinem damaligen Assistenten, vielfach über die betreffenden Fragen diskutierte, aber doch in voller Selbständigkeit des talentvollen Autors. Ich selbst bin nicht mit allem einverstanden, was der Autor mitteilt, da nun aber der Tod jede Aussprache verhindert hat, so glaube ich, am zweckmäßigsten zu handeln, wenn ich die Arbeiten ohne jede Änderung hier drucken lasse. — Bei den Korrekturen bin ich von der Witwe, Frau Dr. Zoeppritz, und von Frl. Dr. Borchers unterstützt worden; ich möchte dafür auch an dieser Stelle besten Dank sagen.

VIIa. Nachtrag zu den Laufzeitkurven, von Karl Zoeppritz (†).

Als großer Mangel der aufgestellten Laufzeitkurven mußte gelten, daß sie nicht weiter als bis 13000 km reichten. Ja, das letzte Stück von 10000 bis 13000 mußte schon als recht unsicher

gelten, da hierfür nur ein einziger Beobachtungspunkt, nämlich der Einsatz des San Franzisko-Erdbebens in Tiflis vorhanden war und zwar in 11100 km Entfernung. Darüber hinaus stellte die Laufzeitkurve eine vorläufige Annahme dar, nämlich wie die Kurve verlaufen müßte, wenn die Fortpflanzungsgeschwindigkeit im Kern bis zu einer Tiefe von etwa 3000 km von der Erdoberfläche aus gerechnet konstant bliebe. Ich erwartete deshalb mit großer Spannung das Eintreffen des Dezemberberichts 1906 der seismischen Station in Apia auf Samoa. Dr. Linke hatte schon brieflich Mitteilung gemacht über das große Beben vom 19. Dez. 1906, das in Samoa direkt gefühlt wurde, viel stärker jedoch noch auf der zur Tongagruppe gehörenden Insel Vavau. Beginn und Epizentralpunkt dieses Bebens können mit einer für die vorläufige Entscheidung genügenden Genauigkeit festgelegt werden. Trotzdem möchte man wohl ungern auf ein einziges Erdbeben die Entscheidung in einer solchen Frage gründen. Da trifft es sich nun günstig, daß noch zwei andere Beben herangezogen werden können, das vom Vortage, nämlich dem 18. Dez., und das sehr kräftige Beben vom 2. Jan. 1907, deren Herde nicht weit entfernt waren von dem des Erdbebens am 19. Dez. 1906. Alle drei Beben scheinen nämlich nach Dr. Linke und Dr. Angenheister im Zusammenhang zu stehen mit dem Steilabsturz östlich der Linie Samoa-Inseln—Tonga-Inseln, der bis 7 km Meerestiefe führt.

Die Daten der drei Beben ergeben folgendes:

					$L-P$
18. XII. 06	Samoa	P:	$21^h\ 0^m 50^s$	?	?
19. XII. 06	„	eP:	$1^h 16^m\ 0^s$	iL: $17^m\ 1^s$	61^s
2. I. 07	„	iP:	$11^h 57^m 27^s$	$(L?)$: $58^m 37^s$	70^s

Danach kann man die Entfernung des Epizentralpunktes am 19. XII. auf etwa 600 am 2. I. auf 700 km schätzen und wird auch beim Beben vom 18. XII. nicht weit fehlgehen, wenn man hier die Entfernung ebenfalls zu 600 km annimmt. Es sind somit, um die Zeit des Bebens im Epizentralpunkt zu erhalten, von dem ersten Einsatz in Samoa zu subtrahieren die Zeiten $1^m 26^s$, bezw. $1^m 26^s$, bezw. $1^m 40^s$. Darnach erhalten wir für die Zeiten der Erdbeben

18. XII. 06: $20^h 59^m 24^s$
19. XII. 06: $1^h 14^m 34^s$
2. I. 07: $11^h 55^m 47^s$

Die Einsätze in Göttingen sind dann die folgenden:

18. XII. 06	*P*: $21^h 18^m 59^s$	Laufzeit 1175	*S*:		Laufzeit
19. XII. 06	*P*: $1^h 34^m 12^s$	„ 1178	„	$1^h 47^m 48^s$	2994
2. I. 07	*P*: $12^h 15^m 46^s$	„ 1199	„	$12^h 29^m 31^s$	2025

Für die Entfernung von Göttingen können dabei gelten die Zahlen 16.5, 16.5 und 16.6 Megameter.

Vergleiche ich diese Laufzeiten mit den bis 13 Megam. reichenden Laufzeitkurven, so ergibt sich, daß die Punkte sowohl für die Longitudinal- wie für die Transversalwellen und zwar besonders für die ersteren erheblich über den Kurven liegen, wie man sie erhält unter Annahme konstanter Geschwindigkeit im Kern. Sollten etwa die ersten im Vertikalseismometer besonders kräftigen Einsätze in Göttingen garnicht den direkt durchs Erdinnere gegangenen Wellen entsprechen? Alle sich hier ergebenden Möglichkeiten wurden von Prof. Wiechert diskutiert, besonders eingehend der Fall einer Reflexion an der Grenze des Kerns. Doch scheint es nach allem zur Zeit ziemlich sicher, daß wir in den oben angegebenen Welleneinsätzen die direkten durchs Erdinnere zu uns gelangten Wellen vor uns haben. Daraus folgt aber mit Sicherheit, wenn die Laufzeitkurve bis etwa 5000 km richtig ist, daß in tieferen Schichten der Erde eine Abnahme der Geschwindigkeit besonders der Dilatationswellen eintreten muß. In welcher Weise diese Abnahme erfolgt, wird sich freilich mit dem heutigen Beobachtungsmaterial noch nicht einwandfrei entscheiden lassen. Ich habe vorläufig die Rechnung einmal durchgeführt unter der Annahme, daß die Strahlkurven wie im Mantel so auch im Kern in erster Annäherung Kreisbögen sind oder daß, was dasselbe bedeutet, die Geschwindigkeit im Kern abnimmt nach einer Funktion parabolischen Charakters. Dann erhebt sich freilich die entsprechende Laufzeitkurve schon von etwa $\Delta = 5$ Megametern ab etwas über die in Tafel I angegebene. Doch beträgt der Unterschied bei 8000 km erst etwa 15^s. Es wären dann für den Einsatz des kalifornischen Bebens noch die Stationen: Samoa, Upsala, Straßburg zu berücksichtigen, zu denen nach dem Jahresbericht von 1906 auch noch Potsdam kommt. Leider ist der Einsatz des kalifornischen Bebens so wenig scharf, daß man keine Gewißheit hat, ob die Angaben aller Stationen sich auf dieselbe Welle beziehen. Sollte dagegen die Laufzeitkurve bis etwa 10 Megameter tatsächlich den auf Tafel I verzeichneten Verlauf haben, so müßte die Geschwindigkeitsabnahme erst in größerer Tiefe eintreten, aber dann entsprechend schneller vor sich gehen. Die Approximationsrechnung ließe sich dann in der

Weise durchführen, daß man eine äußere Schicht mit schneller Zunahme der Geschwindigkeit, eine mittlere mit konstanter Geschwindigkeit und eine innere mit parabolisch abnehmender Geschwindigkeit zu Grunde legte. Die Entscheidung werden die Erdbebenbeobachtungen der nächsten Jahre bringen. Jede Abnahme der Geschwindigkeit im Erdinnern bedeutet für die Oberfläche, daß einer gewissen Zone weniger Energie durch die betreffende Wellenart zugeführt wird. Erfolgt die Geschwindigkeitsabnahme mit der Tiefe sehr schnell, so wird, wie Benndorf gezeigt hat, die betreffende Zone ganz von den betreffenden Vorläufern entblößt. Da eine Geschwindigkeitsabnahme für die ersten Vorläufer schon heute als erwiesen gelten kann, so richtete ich mein Augenmerk besonders auch darauf, ob es eine Entfernung gibt, aus der die ersten Vorläufer auch im Diagramm des Vertikalseismometers auffallend schwach oder garnicht hervortreten. Auch hier liegt die endgültige Entscheidung erst in der Zukunft. Doch scheint es fast, als träte ein solcher Vorläuferschatten tatsächlich für die Dilatationswellen in $\varDelta = 13-14$ Megam. auf. Ich stütze mich dabei auf zwei Erdbeben, die in Deutschneuguinea gefühlt wurden am 14. IX. 06 und am 2. X. 06. Ziehe ich wieder den Samoabericht heran, so haben wir folgendes:

14. IX. 06 Samoa P: $16^h 12^m 11^s$ S: $16^h 18^m 33^s$ $S-P$: 382^s
2. X. 06 „ P: $1^h 59^m 11^s$ S: $2^h 5^m 36^s$ $S-P$: 385^s

Die Zeitdifferenz $S-P$ zwischen ersten und zweiten Vorläufern läßt nach unserer Laufzeitkurve Tafel I auf eine Entfernung von 4700—4800 km schließen, was bis auf etwa 100 km übereinstimmt mit der Entfernung von Apia nach Finschhafen, wo das erste Beben gefühlt wurde. Daraus berechnet sich weiter als Zeit des ersten Stoßes, indem vom ersten Einsatz in Samoa $8^m 16^s$ abgezogen werden, am 14. IX. 06: $16^h 3^m 55^s$ und am 2. X. 06: $1^h 50^m 55^s$. Vergleiche ich damit die Eintreffzeiten der Dilatationswellen in Göttingen, so haben wir bei beiden Beben zunächst einen sehr schwachen Einsatz und zwar am 14. IX. 06 um $16^h 23^m 40^s$, am 2. X. 06 um $2^h 10^m 34^s$, was die Laufzeiten 1185 u. 1179 Sek. ergibt. Etwa 2 Minuten später trifft bei beiden Erdbeben eine stärkere Wellengruppe ein, nämlich um $16^h 25^m 14^s$ bez. um $2^h 12^m 22^s$, dies ergibt die bezügl. Laufzeiten 1281 Sek. und 1287 Sek.. War der Herd des Bebens nur 4600 km von Samoa entfernt, so erniedrigen sich die Laufzeiten dieses zweiten Einsatzes auf 1271 und 1277 Sek. Da die Entfernung des Herdes von Göttingen etwa 13600 km beträgt, so dürfte dieser zweite Einsatz mit großer Wahrscheinlich-

keit den einmal an der Erdoberfläche reflektierten Wellen entsprechen, indem wir für die halbe Wegstrecke $\varDelta = 6800$ eine Laufzeit von 637 Sek. annehmen.

Die beiden Neuguinea-Erdbeben würden uns demnach folgendes lehren: Der Einsatz der direkten Dilatationswellen ist in 13600 km Entfernung so schwach, daß er wenigstens am 14. IX. 06 und am 2. X. 06 in Göttingen auch vom Vertikalseismometer nicht mehr registriert wurde. Nach 1170—1180 Sek. taucht eine sehr schwache Wellengruppe auf, deren Ursprung noch nicht mit Bestimmtheit angegeben werden kann. Nach 1270—1280 Sek. setzen weit kräftiger dann die Dilatationswellen ein, die einmal in 6800 km. Entfernung an der Erdoberfläche reflektiert worden sind. Die zweimal an der Erdoberfläche reflektierten Dilatationswellen heben sich so wenig hervor, daß ihr Einsatz nicht mit einiger Sicherheit abgelesen werden kann.

Wie verhalten sich nun die Scherungswellen bei einem Beben von 13600 km Entfernung? Auch darüber geben uns die Erdbeben vom 14. IX. und 2. X. 06 einigen Aufschluß. Bei beiden Beben ist übereinstimmend der erste Einsatz der zweiten Vorläufer so schwach, daß er nicht mit Sicherheit sich angeben läßt. Die Region der Wechselwellen ist durch unregelmäßige Bewegungen gekennzeichnet. Sehr deutlich treten dann in den Diagrammen beider Beben die Einsätze der einmal und noch stärker die der zweimal (müßte sehr scharf transversal sein) an der Erdoberfläche reflektierten Wellen hervor. Die Zeiten sind für Göttingen die folgenden:

4. IX. 06	SR_1:	$41^m 42^s$;	also Laufzeit	$2267 = 2 \times 1133$ Sek.	entspricht
2. X. 06	„	$28^m 56^s$;	„ „	$2281 = 2 \times 1140$ Sek.	2×6800 km
4. IX. 06	SR_2:	$46^m 30^s$;	„ „	$2555 = 3 \times 852$ Sek.	entspricht
2. X. 06	SR_2:	$34^m\ 5^s$;	„ „	$2590 = 3 \times 863$ Sek.	3×4530 km

Schließlich möchte ich noch auf eine letzte Wellengruppe hinweisen, der beim Erdbeben vom 14. IX. eine Laufzeit von etwa 3030 Sek., beim Beben vom 2. X. 06 eine solche von etwa 3050 Sek. zukommt. Diese Wellengruppe hat auffallend unregelmäßigen Charakter und entspricht möglicherweise Scherungswellen, die sehr oft an der Erdoberfläche reflektiert worden sind. Aus den Formeln für die Reflexion von Erdbebenwellen folgt nämlich, daß eine sehr flach einfallende linear polarisierte Transversalwelle, deren Schwingungen in der Einfallsebene erfolgen, bei der Reflexion nichts von ihrer Energie als Longitudinalwelle weiterschickt. Da nun linear po-

larisierte Transversalwellen mit Schwingungen senkrecht zur Einfallsebene überhaupt keinen Verlust durch weitergesandte Dilatationswellen erleiden, so wäre es wohl denkbar, daß bei einem sehr fernen Beben die 3, 4 u. s. w. mal reflektierten Transversalwellen im Diagramm sich nicht mehr abheben, wohl aber wieder die noch öfter reflektierten, die wegen ihres flachen Einfalls nur sehr wenig Energie an die Longitudinalwellen abgegeben haben. Sie treffen nahezu gleichzeitig ein und werden deshalb in komplizierter Weise miteinander interferieren, womit auch der unregelmäßige, von der Sinusform stark abweichende Charakter dieser Wellengruppe verständlich wäre.

Der Beginn der sogenannten Hauptwellen, die wenigstens zu einem Teil jedenfalls den von Rayleigh aus der Theorie gefolgerten Oberflächenwellen entsprechen, tritt besonders deutlich meist im Vertikalseismometer hervor, während in den Horizontalkomponenten, wie es scheint, die Überlagerung durch andere Wellenarten im Beginn der Hauptwellen oft störend wirkt. Bei den beiden betrachteten Neuguineabeben finde ich die Laufzeit der ersten Oberflächenwellen mit großer Periode (über $^1/_2$ Minute) zu etwa 50 bis 51 Min., was einer Fortpflanzungsgeschwindigkeit von etwa 4,5 km/sec. entspäche.

Was nun die Geschwindigkeitsverteilung in Abhängigkeit von der Entfernung vom Erdmittelpunkt anlangt, so liefert die auf Tafel [1]) angegebene Geschwindigkeitskurve die auf Tafel [1]) gezeichnete Laufzeitkurve. Da vorläufig eine genauere Festlegung der Laufzeitkurve nicht möglich war, so nahm ich für die Rechnung zunächst an, die Kurve sei für 9000 km etwa 20 Sek. zu tief angenommen, sie erhöhe sich also von etwa 5000 ab allmählich über die gezeichnete, um in stetigem Verlauf durch den durch die beiden Tongaerdbeben festgelegten Punkt: $\Delta = 16000$ km, $T = 1150$ Sek. hindurch zu gehen. Eine solche Laufzeitkurve erhält man bei der einfachen Annahme, daß die Strahlkurven der Longitudinalwellen sich durch 2 Kreisbögen approximieren lassen, das heißt, daß die Erde in eine innere Kugel und eine konzentrische äußere Schale zerfällt, daß ferner in der äußeren (Gestein-) Schale die Geschwindigkeit der Dilatationswellen nach einem parabolischen Gesetz zu-, in der Kernkugel dagegen nach einem analogen Gesetz abnimmt. Rechnen wir mit diesen Annahmen, so haben wir die auf nächster Seite folgende Tabelle, die bis $\Delta = 16$ Megameter reicht.

1) Vergl. die Anmerkung am Schluß der Arbeit.

Vergleicht man die unter den gemachten Annahmen erhaltene Laufzeitkurve mit der früheren, so sieht man sofort, wie gering

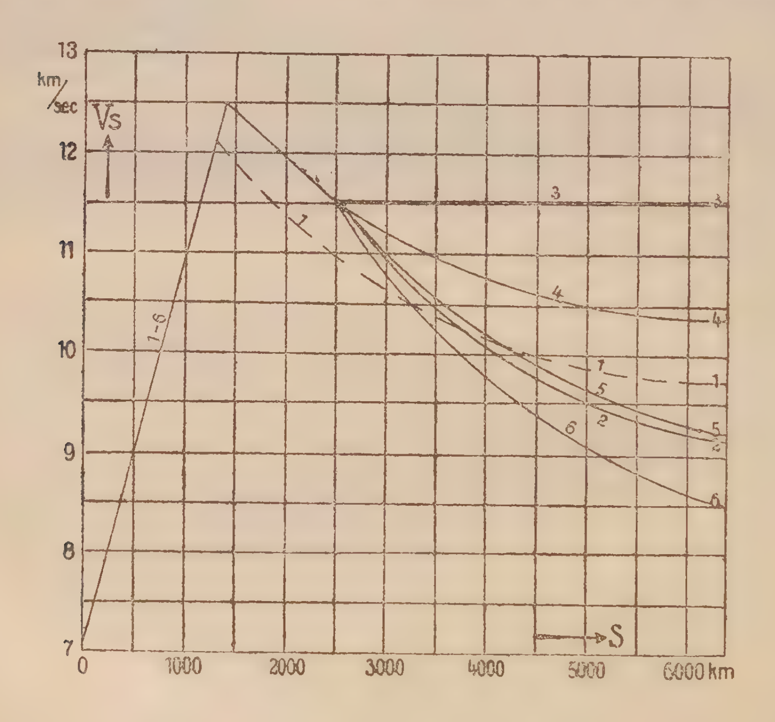
Entfernung in Megam.	sin i	Laufzeit [1])
0		
1		
2		
3		
4		
5	0,4658	
6	465	
7	462	
8	459	703
9	453	
10	445	828
11	433	
12	415	
13	393	
14	363	
15	329	
16	290	1150

die Unterschiede beider Kurven bis zu $\varDelta = 9000$ km nur sind. Dieser Vergleich allein lehrt schon, wie genau in diesem Bereich die Laufzeitkurve bekannt sein muß, wenn man einigermaßen sichere Schlüsse auf die spezielle Geschwindigkeitsverteilung im Erdinnern ziehen will. Daß eine Geschwindigkeitsabnahme als erwiesen gelten muß, habe ich schon ausgeführt. Alles weitere ist aber vorläufig nur als Rechnungsresultat zu betrachten. Mit den zuletzt gemachten Annahmen und der aus ihnen zu folgernden Laufzeitkurve (Tabelle [1]) haben wir den in Kurve [1]) Tafel [1]) verzeichneten Geschwindigkeitsverlauf im Erdinnern. Eine ganz ähnliche Laufzeitkurve, der nach den heutigen Beobachtungsdaten etwa die gleiche Berechtigung zukäme, würde sich aber auch ergeben, wenn wir eine ganz erheblich abweichende Geschwindigkeitsverteilung im Innern der Erde annehmen, etwa von der Erdoberfläche aus nach dem ersten starken Anstieg eine Schale konstanter Geschwindigkeit und dann erst eine Schale mit entsprechend stärkerer Abnahme der Geschwindigkeit. Eine gewisse Unterstützung der Forschung in dieser Richtung dürfen wir erhoffen, wenn gewisse Entfernungszonen auf der Erdoberfläche gefunden würden, die von den direkt durchs Erdinnere laufenden Wellen nicht betroffen werden. Aus den etwa aufgefundenen Schattenregionen kann man sich dann vielleicht ein sichereres Bild machen über die Gebiete abnehmender

1) Vergleiche die Anmerkung am Schluß der Arbeit.

Geschwindigkeit im Erdinnern, als aus der Laufzeitkurve allein. Wäre die Laufzeitkurve mit beliebiger Genauigkeit bekannt, so genügte sie zwar schon allein, um das Problem der Geschwindigkeitsverteilung zu lösen. Da wir aber von einer auf die Sekunde genauen Kenntnis der Laufzeiten noch weit entfernt sind, so müssen wir unser Augenmerk auch auf andere Hülfsmittel richten. In diesem Sinne sind eben die Beobachtungen an den beiden Neuguinea-Erdbeben zu verstehen, wonach es in der Tat scheint, daß wir uns in 13 600 km Entfernung in einer Region des Schattens oder wenigstens stark verminderter Energie für die direkt laufenden Dilatationswellen befinden.

Anmerkung von E. Wiechert. Die Tafeln, welche der Verfasser auf Seite 62 und 63 zitiert, fehlen, die Zahlen der Tabelle auf Seite 63 sind unvollständig. Offenbar aber sind Untersuchungen gemeint, auf welche der Verfasser in jener nachgelassenen Arbeit bezug nimmt, deren Resultate im Jahrgang 1912 dieser „Nachrichten“ S. 121—206 veröffentlicht wurden. Die in Betracht kommenden Zahlen mit der zugehörigen Figur setze ich hier nochmals her. In der Figur handelt es sich um die Kurve 1.



Laufzeiten für die 1. Vorläufer.

Entfernung in 10^3 km	Laufzeiten in sec. nach Wiechert-Zoeppritz 1906	nach der neuen Zoeppritz'schen Annahme
1	136	136
2	257	257
3	358	358
4	442	441
5	512	510
6	572	573
7	631	640
8	688	704
9	743	768
10	795	830
11	844	890
12	888	949
13	929	1006
14		1059
15		1107
16		1148

VIIb. Über Reflexion und Durchgang seismischer Wellen durch Unstetigkeitsflächen.

Von

Karl Zoeppritz (†).

Vorgelegt in der Sitzung vom 27. Juni 1914 von E. Wiechert.

Professor E. Wiechert hat in einer kürzlich erschienenen Arbeit (E. Wiechert und K. Zoeppritz: Über Erdbebenwellen.) die Reflexion seismischer Wellen an der Erdoberfläche behandelt. Dabei wurden Luft- und Wasserbedeckungen wegen ihres verschwindenden Einflusses vernachlässigt und die Reflexion so behandelt, als ob sie am leeren Raume erfolgte. Trotzdem diese Annahme bei der heute noch unvermeidlichen Ungenauigkeit der seismischen Meßmethoden, wie eine rohe Schätzung lehrt, durchaus gerechtfertigt ist, so scheint es doch nicht ohne Interesse, die Formeln aufzustellen für die Reflexion und den Durchgang elastischer Wellen durch die Grenzfläche zwischen zwei beliebig gearteten elastischen Medien. Und dies nicht nur im Hinblick auf die Verhältnisse an der Erdoberfläche, sondern vor allem auch wegen der möglichen Anwendung auf etwaige Grenzflächen im Innern der Erde. Ich habe deshalb auf Anregung von Professor Wiechert diesen allgemeineren Fall behandelt und werde im Folgenden einige Resultate angeben.

Ich betrachte die ebene Grenzfläche zwischen zwei Medien 1 und 2 mit den Dichten ϱ_1 und ϱ_2 und den elastischen Konstanten $\mathfrak{a}_1$, $\mathfrak{b}_1$ und $\mathfrak{a}_2$, $\mathfrak{b}_2$. $\mathfrak{a}$ und $\mathfrak{b}$ bedeuten dabei die Fortpflanzungsgeschwindigkeiten der reinen Dilatations- und der dilatationsfreien Scherungswellen. Zur leichteren Übersicht führe ich kurz die Formeln an, die den Zusammenhang mit den Lamé'schen Konstanten

λ und μ und mit der Poisson'schen Konstante σ angeben

$$\frac{\lambda + 2\mu}{\varrho} = \mathfrak{a}^2, \quad \frac{\mu}{\varrho} = \mathfrak{b}^2, \quad \sigma = \frac{\mathfrak{a}^2 - 2\mathfrak{b}^2}{2(\mathfrak{a}^2 - \mathfrak{b}^2)}.$$

Bezüglich der Bewegungsgleichungen für isotrope, elastische Medien kann ich auf die betreffenden Ausführungen von E. Wiechert in der schon zitierten Arbeit verweisen. Zur leichteren Orientierung möchte ich sie hier noch kurz für eine Komponente in der von englischen Autoren meist gebrauchten Form [1]) anschreiben:

$$\varrho(\mathfrak{a}^2 - \mathfrak{b}^2)\frac{\partial \Delta}{\partial x} + \varrho . \mathfrak{b}^2 \nabla^2 u = \varrho \frac{\partial^2 u}{\partial t^2},$$

worin Δ die Dilatation $\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)$

und ∇^2 den Laplaceschen Operator $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ bedeuten.

Während bei der Reflexion am leeren Raum nur die 2 Konstanten $\mathfrak{a}$ und $\mathfrak{b}$ auftraten, haben wir in unserm allgemeineren Fall deren 6, dürfen also keineswegs so einfache Formeln erwarten wie in jenem Spezialfall. Ich beschränke mich auf den Fall einer periodischen Wellenbewegung. Die erhaltenen Formeln werden wenigstens qualitativ das Verhalten auch bei stoßartigen, unperiodischen Wellen darstellen. Genau wie bei der Reflexion am leeren Raum gelten gewisse geometrische Beziehungen zwischen den Winkeln der einfallenden bezw. der reflektierten und der durchgehenden Wellenstrahlen, die allein durch die Verhältnisse der Geschwindigkeiten bestimmt werden. Stets denke ich mir, daß die Welle aus dem Medium 1 kommend auf die ebene Grenzfläche 1, 2 fällt. Es sind 2 Fälle möglich: die einfallende Welle kann entweder longitudinalen oder sie kann transversalen Charakter haben. Im ersteren Fall haben wir es mit reinen Dilatationswellen, im letzteren umgekehrt mit dilatationsfreien, reinen Scherungswellen zu tun. Die transversalen Wellen denke ich mir noch weiter zerlegt in zwei Komponenten, von denen die eine in der Einfallsebene, die andere senkrecht dazu schwingt. So haben wir im Ganzen drei Fälle zu untersuchen, je nach Art der einfallenden Welle. Der leichteren Übersicht halber bezeichne ich die spitzen Winkel zwischen Strahlrichtung und Einfallslot mit zwei Buchstaben. Der erste von diesen, e, r oder d, deutet an, daß der einfallende, reflektierte oder durchgehende Strahl gemeint ist, der zweite

1) Vergl. A. E. H. Lowe, Lehrbuch der Elastizität, Deutsch von A. Timpe. 1907.

Buchstabe, l oder t, gibt an, ob die Welle longitudinal oder transversal ist. Es gelten dann nach dem Brechungsgesetz die Beziehungen:

$$\sin el : \sin et : \sin rl : \sin rt : \sin dl : \sin dt = \mathfrak{a}_1 : \mathfrak{b}_1 : \mathfrak{a}_1 : \mathfrak{b}_1 : \mathfrak{a}_2 : \mathfrak{b}_2$$

Das Koordinatensystem lege ich so, daß die z-Axe vom Medium 1 nach dem Medium 2, in der Zeichnung von unten nach oben, weist, die x-Axe geht in der Einfallsebene der Grenzfläche parallel von links nach rechts. Die y-Axe verläuft senkrecht dazu von vorn nach hinten. Die Verrückungen nenne ich u, v, w, entsprechend den Axen x, y, z, und zwar u_1, v_1, w_1 für das Medium 1, u_2, v_2, w_2 für das Medium 2; ein wagerechter Strich über dem Buchstaben bedeutet, daß der betreffende Wert für die Grenzfläche 1, 2 gemeint ist.

Für die Grenzfläche sind folgende Grenzbedingungen zu erfüllen: 1) muß die Summe der Horizontal- bezw. der Vertikalverrückungen im Medium 1 gleich der entsprechenden Summe im Medium 2 sein. In Formeln:

$$1)\ \Sigma\bar{u}_1 = \Sigma\bar{u}_2, \qquad 2)\ \Sigma\bar{v}_1 = \Sigma\bar{v}_2, \qquad 3)\ \Sigma\bar{w}_1 = \Sigma\bar{w}_2$$

2) muß die Summe der Drucke an der Grenzfläche im Medium 1 gleich der entsprechenden im Medium 2 sein. Also muß erfüllt sein:

$$4) \qquad \sum \varrho_1 \mathfrak{b}_1^2 \overline{\left(\frac{\partial u_1}{\partial z} + \frac{\partial w_1}{\partial x}\right)} = \sum \varrho_2 \mathfrak{b}_2^2 \overline{\left(\frac{\partial u_2}{\partial z} + \frac{\partial w_2}{\partial x}\right)}$$

$$5) \qquad \sum \varrho_1 \mathfrak{b}_1^2 \overline{\left(\frac{\partial v_1}{\partial z} + \frac{\partial w_1}{\partial y}\right)} = \sum \varrho_2 \mathfrak{b}_2^2 \overline{\left(\frac{\partial v_2}{\partial z} + \frac{\partial w_2}{\partial y}\right)}$$

$$6) \qquad \sum\left\{\varrho_1 \mathfrak{a}_1^2 \overline{\Delta}_1 - 2\varrho_1 \mathfrak{b}_1^2 \overline{\left(\frac{\partial u_1}{\partial x} + \frac{\partial v_1}{\partial y}\right)}\right\}$$
$$= \sum\left\{\varrho_2 \mathfrak{a}_2^2 \overline{\Delta}_2 - 2\varrho_2 \mathfrak{b}_2^2 \overline{\left(\frac{\partial u_2}{\partial x} + \frac{\partial v_2}{\partial y}\right)}\right\}.$$

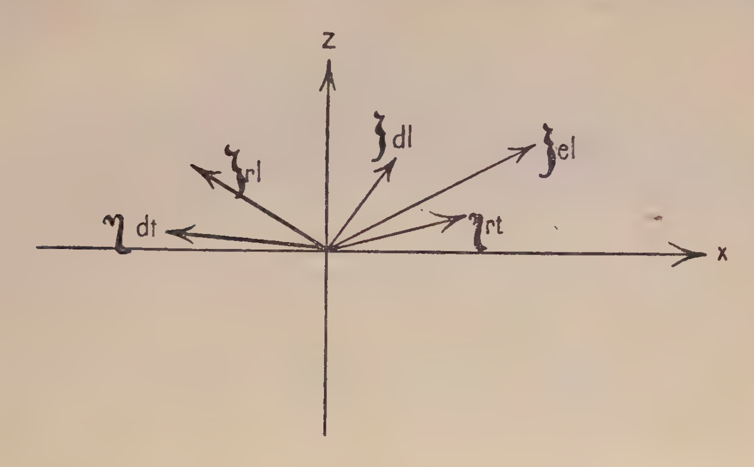
Hierin bedeutet Δ die Dilatation $\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)$ und $\overline{\Delta}$ den Wert der Dilatation für die Grenze.

1. Einfallende Longitudinalwelle.

Es treten nur Verrückungen in der Einfallsebene auf. Bezeichne ich der Reihe nach für die einfallende, die reflektierte und die durchgehende Longitudinalwelle die Verrückung mit ξ_{el}, ξ_{rl}, ξ_{dl}, entsprechend für die Transversalwelle mit η_{et}[1]), η_{rt}, η_{dt},

1) Tritt in unserm Fall nicht auf.

wobei die als positiv angenommenen Richtungen durch die Pfeile in der folgenden Figur gekennzeichnet sind, so können wir für die



ebenen Wellen folgende Ansätze machen, von denen jeder einzelne die für das Medium gültige Wellengleichung erfüllt.

$$\xi_{el} = \mathfrak{R}\left\{M_{el} \cdot e^{\frac{2\pi i}{T}\left(t - \frac{\sin el\, x + \cos el\, z}{\mathfrak{a}_1}\right)}\right\} \quad \begin{aligned} u &= \xi_{el} \sin el \\ v &= 0 \\ w &= \xi_{el} \cos el \end{aligned}$$

$$\xi_{rl} = \mathfrak{R}\left\{M_{rl} \cdot e^{\frac{2\pi i}{T}\left(t - \frac{\sin el\, x - \cos el\, z}{\mathfrak{a}_1}\right)}\right\} \quad \begin{aligned} u &= -\xi_{rl} \sin el \\ v &= 0 \\ w &= \xi_{rl} \cos el \end{aligned}$$

$$\xi_{dl} = \mathfrak{R}\left\{M_{dl} \cdot e^{\frac{2\pi i}{T}\left(t - \frac{\sin dl\, x + \cos dl\, z}{\mathfrak{a}_2}\right)}\right\} \quad \begin{aligned} u &= \xi_{dl} \sin dl \\ v &= 0 \\ w &= \xi_{dl} \cos dl \end{aligned}$$

$$\eta_{et} = \mathfrak{R}\left\{0 \cdot e^{\frac{2\pi i}{T}\left(t - \frac{\sin et\, x + \cos et\, z}{\mathfrak{b}_1}\right)}\right\} \quad \begin{aligned} u &= 0 \\ v &= 0 \\ w &= 0 \end{aligned}$$

$$\eta_{rt} = \mathfrak{R}\left\{M_{rt} \cdot e^{\frac{2\pi i}{T}\left(t - \frac{\sin et\, x - \cos et\, z}{\mathfrak{b}_1}\right)}\right\} \quad \begin{aligned} u &= \eta_{rt} \cdot \cos et \\ v &= 0 \\ w &= \eta_{rt} \cdot \sin et \end{aligned}$$

$$\eta_{dt} = \mathfrak{R}\left\{M_{dt} \cdot e^{\frac{2\pi i}{T}\left(t - \frac{\sin dt\, x + \cos dt\, z}{\mathfrak{b}_2}\right)}\right\} \quad \begin{aligned} u &= -\eta_{dt} \cdot \cos dt \\ v &= 0 \\ w &= \eta_{dt} \sin dt \end{aligned}$$

Geht man mit diesem Ansatz in die Grenzbedingungen 1 bis 6 ein, von denen die Bedingungen 2 und 5 für unsern Fall wegfallen, so hat man 4 Gleichungen, die linear und homogen sind in den 5 Größen M_{el}, M_{rl}, M_{dl}, M_{rt}, M_{dt}. Nehme ich die Amplitude M_{el} der einfallenden Welle als bekannt an, so sind also die andern 4 Größen durch die 4 Grenzbedingungen bestimmt. Zur Bestimmung der Verhältnisse der 5 Größen M_{el}, M_{rl}, M_{dl}, M_{rt}, M_{dt} erhält man für die Faktoren der Gleichungen folgende Matrix:

M_{el}	M_{rl}	M_{rt}	M_{dl}	M_{dt}
$+\sin el$	$-\sin el$	$+\cos et$	$-\sin dl$	$+\cos dt$
$+\cos el$	$+\cos el$	$+\sin et$	$-\cos dl$	$-\sin dt$
$-\sin 2el$	$-\sin 2el$	$+\frac{\mathfrak{a}_1}{\mathfrak{b}_1}\cos 2et$	$+\frac{\varrho_2}{\varrho_1}\left(\frac{\mathfrak{b}_2}{\mathfrak{b}_1}\right)^2\left(\frac{\mathfrak{a}_1}{\mathfrak{a}_2}\right)\sin 2dl$	$-\frac{\varrho_2}{\varrho_1}\left(\frac{\mathfrak{b}_2}{\mathfrak{b}_1}\right)^2\left(\frac{\mathfrak{a}_1}{\mathfrak{b}_2}\right)\cos 2d$
$-\cos 2et$	$+\cos 2et$	$+\frac{\mathfrak{b}_1}{\mathfrak{a}_1}\sin 2et$	$+\frac{\varrho_2}{\varrho_1}\frac{\mathfrak{a}_2}{\mathfrak{a}_1}\cos 2dt$	$+\frac{\varrho_2}{\varrho_1}\frac{\mathfrak{b}_2}{\mathfrak{a}_1}\sin 2dt.$

Die Formeln, die M_{rl}, M_{rt}, M_{dl}, M_{dt} mit M_{el} verbinden, lassen sich auf Grund dieser Matrix ohne weiteres hinschreiben, erlauben auch noch einige Vereinfachung. Doch geht diese Vereinfachung nicht sehr weit, da wir es, wie schon bemerkt, mit 6 völlig von einander unabhängigen Materialkonstanten zu tun haben. Erst wenn man über die Natur dieser Konstanten gewisse vereinfachende Annahmen macht, lassen sich die Formeln bedeutend abkürzen. Setzen wir z. B. $\varrho_2 = 0$, so haben wir den Fall der Reflexion am leeren Raum, es fallen die obersten beiden Zeilen und die letzten beiden Kolumnen unserer Matrix fort, und man gelangt zu den von Prof. Wiechert abgeleiteten Formeln. Ist nur $\mathfrak{b}_2 = 0$, d. h. hat das Medium 2 verschwindende Steifigkeit (Reflexion an Flüssigkeiten oder Gasen), so lautet unsere Matrix:

M_{el}	M_{rl}	M_{rt}	M_{dl}	M_{dt}
$+\sin el$	$-\sin el$	$+\cos et$	$-\sin dl$	1
$+\cos el$	$+\cos el$	$+\sin et$	$-\cos dl$	0
$-\sin 2el$	$-\sin 2el$	$+\frac{\mathfrak{a}_1}{\mathfrak{b}_1}\cos 2et$	0	0
$-\cos 2et$	$+\cos 2et$	$+\frac{\mathfrak{b}_1}{\mathfrak{a}_1}\sin 2et$	$+\frac{\varrho_2}{\varrho_1}\frac{\mathfrak{a}_2}{\mathfrak{a}_1}$	0

Setzt man die Amplitude der einfallenden Welle $M_{el} = 1$, so leitet man aus dieser Matrix leicht die Formeln ab:

$$M_{rl} = -\frac{1 - f(1-m)}{1 + f(1+m)}$$

$$M_{rt} = \frac{4 \frac{\mathfrak{b}_1 \varrho_1}{\mathfrak{a}_2 \varrho_2} \sin el \,.\cos dl \,.\cos 2\,et}{1 + f(1+m)}$$

$$M_{dl} = \frac{2 \frac{\mathfrak{a}_1 \varrho_1}{\mathfrak{a}_2 \cdot \varrho_2} \cos 2\,et}{1 + f(1+m)}.$$

Hier bedeuten f und m folgende Abkürzungen:

$$f = \frac{\mathfrak{a}_1 \varrho_1}{\mathfrak{a}_2 \varrho_2} \frac{\cos dl \cos^2 2\,et}{\cos el}$$

$$m = \left(\frac{\mathfrak{b}_1}{\mathfrak{a}_1}\right)^2 \frac{\sin 2\,el \,.\operatorname{tg} 2\,et}{\cos 2\,et}.$$

Diese Formeln sind geeignet, uns ein Bild zu verschaffen, in welcher Weise eine Luft- oder Wasserbedeckung der Oberfläche die Reflexion der aus der Tiefe auftauchenden Longitudinalwellen beeinflußt. Die folgende kleine Tabelle gibt für einige Einfallswinkel die Amplituden M_{rl} und M_{rt} der reflektierten Longitudinal- und Transversalwellen und zwar einmal bei der Reflexion am leeren Raum, andrerseits bei der Reflexion am Meeresgrund. Für die maßgebenden Konstanten sind dabei folgende Werte zu Grunde gelegt: 1. Für das Gestein der Erdoberfläche $\varrho = 2{,}7$, $\mathfrak{a} = 7{,}2^{\text{km}}/_{\text{sec.}}$, $\mathfrak{b} = 4{,}0^{\text{km}}/_{\text{sec.}}$ 2. Für Meerwasser $\varrho = 1$, $\mathfrak{a} = 1{,}43^{\text{km}}/_{\text{sec.}}$ und $\mathfrak{b} = 0^{\text{km}}/_{\text{sec.}}$

Einfallswinkel	M_{rl} Reflexion am		M_{rt} Reflexion am	
	leeren Raum	Meeresgrund	leeren Raum	Meeresgrund
0	1.00	0.87	0	0
20	0.84	0.71	0.71	0.66
45	0.35	0.26	1.10	1.02
70	0.06	0.00	0.85	0.79
90	0.00		0	0

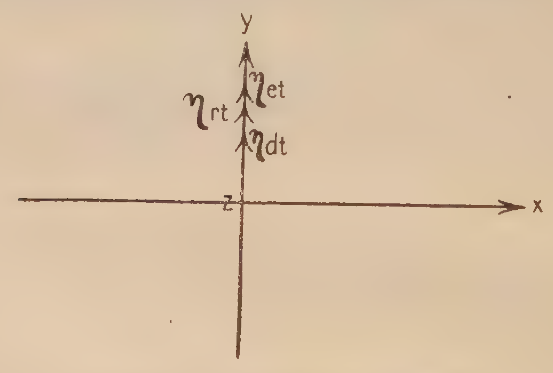
Wie man sieht, sind die Unterschiede zwischen der Reflexion am

Meeresgrund und am leeren Raum nur sehr gering[1]). Sie würden sich bei der angegebenen Genauigkeit von 2 Dezimalen überhaupt nicht mehr bemerkbar machen, wenn wir die Reflexion gegen Luft mit der am leeren Raum verglichen hätten.

Ich habe davon Abstand genommen, die erhaltenen Formeln auf eine Unstetigkeitsfläche im Erdinnern, also etwa auf die Grenzfläche zwischen Gesteinsmantel und Metallkern anzuwenden. Denn einmal ist es keineswegs wahrscheinlich, daß dieser Übergang sehr schnell vor sich geht. Wir kennen im Gegenteil unter den Meteoriten alle Übergangsstufen vom Gesteinsmeteoriten bis zum reinen Eisenmeteorit. Einen ähnlichen Übergang dürfen wir deshalb wohl auch im Innern der Erde vermuten. Andrerseits hätte ein solch rascher Übergang auch in der Laufzeitkurve sich äußern müssen, und zwar vor allem in dem Umstand, daß mit einer sprungartig schnellen Abnahme der Geschwindigkeit in bestimmter Tiefe ein Schattenbereich für die „ersten Vorläufer" an der Erdoberfläche verknüpft ist. Ein solcher hat sich aber bislang nicht nachweisen lassen.

2. Einfallende Transversalwelle.

a) Verrückung senkrecht zur Einfallsebene.



Grenzfläche vom Medium 2 aus gesehen.

1) Damit fällt eine vielleicht ganz überraschende Aussicht, die die Erdbebenforschung hätte eröffnen können. Wäre nämlich der Unterschied der Reflexion am Meeresgrund und am Festland nicht so unbedeutend, so wäre es vielleicht möglich gewesen, aus dem Charakter einer reflektierten Welle zu schließen, ob der Ort der Reflexion Festland oder Meeresgrund ist. Es könnte so die Hoffnung erwachen, mit Hülfe der Erdbebenregistrierung die alte Frage nach der Verteilung von Wasser und Land an den Polen zur Entscheidung zu bringen.

Auch bei der durchgehenden und der reflektierten Welle treten natürlich nur Verrückungen senkrecht zur Einfallsebene auf. Wir machen folgende, der Wellengleichung genügende Ansätze:

$$\begin{cases} \eta_{et} = \Re\left\{M_{et}\,.\,e^{\frac{2\pi i}{T}\left(t-\frac{\sin et\, x + \cos et\, z}{\mathfrak{b}_1}\right)}\right\} & \begin{aligned} u &= 0 \\ v &= +\eta_{et} \\ w &= 0 \end{aligned} \\ \eta_{rt} = \Re\left\{M_{rt}\,.\,e^{\frac{2\pi i}{T}\left(t-\frac{\sin et\, x - \cos et\, z}{\mathfrak{b}_1}\right)}\right\} & \begin{aligned} u &= 0 \\ v &= +\eta_{rt} \\ w &= 0 \end{aligned} \\ \eta_{dt} = \Re\left\{M_{dt}\,.\,e^{\frac{2\pi i}{T}\left(t-\frac{\sin dt\, x + \cos dt\, z}{\mathfrak{b}_2}\right)}\right\} & \begin{aligned} u &= 0 \\ v &= +\eta_{dt} \\ w &= 0 \end{aligned} \end{cases}$$

In diesem Fall liefern die Grenzbedingungen 2) und 5) die nötigen Beziehungen zur Bestimmung der Verhältnisse M_{rt}/M_{et}, M_{dt}/M_{et}. Die Matrix der Koeffizienten lautet hier sehr einfach:

M_{et}	M_{rt}	M_{dt}
1	-1	$-\frac{\varrho_2\,\mathfrak{b}_2 \cos dt}{\varrho_1\,\mathfrak{b}_1 \cos et}$
1	$+1$	$+1$

Setzen wir $M_{et} = 1$, so wird

$$M_{rt} = +\frac{1-k}{1+k}, \quad M_{dt} = \frac{+2}{1+k}, \quad \text{worin } k = \frac{\varrho_2}{\varrho_1}\,\frac{\mathfrak{b}_2}{\mathfrak{b}_1}\,\frac{\cos dt}{\cos et}.$$

b) Verrückungen in der Einfallsebene.

Wieder machen wir gewisse Ansätze, die ebene Wellen darstellen und die Wellengleichung befriedigen. Aus der Figur geht wieder hervor, in welcher Richtung die Verrückungen als positiv gerechnet werden sollen. Bei den transversalen Wellen liegt die Fortpflanzungsrichtung senkrecht zur Richtung der Verrückung η und zwar im Sinn des Uhrzeigers um 90^0 gegen diese gedreht. Die Ansätze für die ebenen Wellen sind hier die folgenden:

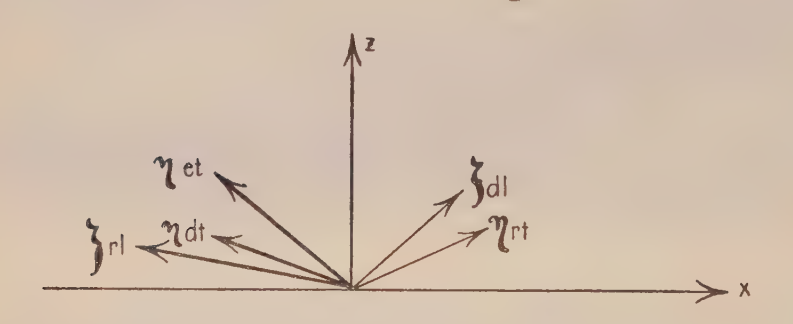
$$\eta_{et} = \Re\left\{M_{et}\,.\,e^{\frac{2\pi i}{T}\left(t-\frac{\sin et\, x + \cos et\, z}{\mathfrak{b}_1}\right)}\right\} \quad \begin{aligned} u &= -\eta_{et}\cos et \\ v &= 0 \\ w &= +\eta_{et}\sin et \end{aligned}$$

$$\eta_{rt} = \Re\left\{M_{rt}\,.\,e^{\frac{2\pi i}{T}\left(t-\frac{\sin et\, x - \cos et\, z}{\mathfrak{b}_1}\right)}\right\} \quad \begin{aligned} u &= +\eta_{rt}\cos et \\ v &= 0 \\ w &= +\eta_{rt}\sin et \end{aligned}$$

$$\xi_{rl} = \Re\left\{M_{rl}\,.\,e^{\frac{2\pi i}{T}\left(t-\frac{\sin el\,x - \cos el\,z}{\mathfrak{a}_1}\right)}\right\}\quad\begin{aligned} u &= -\,\xi_{el}\sin el \\ v &= 0 \\ w &= +\,\xi_{el}\cos el\end{aligned}$$

$$\eta_{dt} = \Re\left\{M_{dt}\,.\,e^{\frac{2\pi i}{T}\left(t-\frac{\sin dt\,x + \cos dt\,z}{\mathfrak{b}_2}\right)}\right\}\quad\begin{aligned} u &= -\,\eta_{dt}\cos dt \\ v &= 0 \\ w &= +\,\eta_{dt}\sin dt\end{aligned}$$

$$\xi_{dl} = \Re\left\{M_{dl}\,.\,e^{\frac{2\pi i}{T}\left(t-\frac{\sin dl\,x + \cos dl\,z}{\mathfrak{a}_2}\right)}\right\}\quad\begin{aligned} u &= +\,\xi_{dl}\sin dl \\ v &= 0 \\ w &= +\,\xi_{dl}\cos dl.\end{aligned}$$



Hier sind wieder die Grenzbedingungen 1, 3, 4 und 6 heranzuziehen und liefern folgende Matrix zur Bestimmung der Amplitudenverhältnisse:

M_{et}	M_{rt}	M_{rl}	M_{dt}	M_{dl}
$-\cos et$	$+\cos et$	$-\sin el$	$+\cos dt$	$-\sin dl$
$+\sin et$	$+\sin et$	$+\cos el$	$-\sin dt$	$-\cos dl$
$+\cos 2\,et$	$+\cos 2\,et$	$-\frac{\mathfrak{b}_1}{\mathfrak{a}_1}\sin 2\,el$	$-\frac{\varrho_2}{\varrho_1}\left(\frac{\mathfrak{b}_2}{\mathfrak{b}_1}\right)\cos 2\,dt$	$+\frac{\varrho_2}{\varrho_1}\frac{\mathfrak{b}_2}{\mathfrak{b}_1}\frac{\mathfrak{b}_2}{\mathfrak{a}_2}\sin 2\,dl$
$-\sin 2\,et$	$+\sin 2\,et$	$+\frac{\mathfrak{a}_1}{\mathfrak{b}_1}\cos 2\,et$	$+\frac{\varrho_2}{\varrho_1}\frac{\mathfrak{b}_2}{\mathfrak{b}_1}\sin 2\,dt$	$+\frac{\varrho_2}{\varrho_1}\frac{\mathfrak{a}_2}{\mathfrak{b}_1}\cos 2\,dt.$

Auch hier werden wie bei der einfallenden Longitudinalwelle die Formeln für den allgemeinen Fall sehr umfangreich, und man wird für die Anwendung vielleicht am bequemsten unmittelbar mit der Matrix rechnen. Ist $\mathfrak{b}_2 = 0$, d. h. hat das Medium 2 verschwindende Steifigkeit, so erhält man folgende Ausdrücke ($M_{et} = 1$ gesetzt):

$$M_{rt} = -\frac{1 - f(1 - m)}{1 + f(1 + m)}$$

$$M_{rl} = +\frac{\cos dl \,.\, \sin 4\,et \,.\, \varrho_1 \,.\, \mathfrak{b}_1}{\cos el \qquad \varrho_2 \,.\, \mathfrak{a}_2} \frac{1}{1 + f(1 + m)}$$

$$M_{dl} = +\frac{2 \sin 2\,et \,.\, \varrho_1 \,.\, \mathfrak{b}_1}{\varrho_2\, \mathfrak{a}_2} \frac{1}{1 + f(1 + m)}$$

$$f = \left(\frac{\mathfrak{b}_1}{\mathfrak{a}_2}\right)^2 \frac{\varrho_1}{\varrho_2} \sin 2\,et \,.\, \sin 2\,dl, \qquad m = \left(\frac{\mathfrak{a}_1}{\mathfrak{b}_1}\right)^2 \frac{\cos 2\,et}{\sin 2\,el \,.\, \operatorname{tg} 2\,et}.$$

Es sei noch für den Fall, wo $\mathfrak{b}_2 = 0$ ist, die Formel für das Verhältnis der Horizontal- zur Vertikalbewegung an der Grenzfläche angeführt:

$$\frac{\Sigma \overline{u}}{\Sigma \overline{w}} = \left(\frac{\mathfrak{a}_1}{\mathfrak{b}_1}\right)^2 \frac{\cos 2\,et}{\sin 2\,el} - \frac{\varrho_2}{\varrho_1} \left(\frac{\mathfrak{a}_2}{\mathfrak{b}_1}\right)^2 \frac{1}{\sin 2\,dl}.$$

Der erste Term gibt das Verhältnis der Horizontalbewegung zur Vertikalbewegung an der Grenzfläche an, also die Tangente des „scheinbaren“ Einfallswinkels, falls kein Medium 2 vorhanden ist. Der zweite Term gibt den Einfluß eines solchen Mediums auf den scheinbaren Einfallswinkel an.

Totalreflexion.

Bisher war in allen Fällen davon abgesehen worden, daß der Sinus eines Winkels zwischen Strahl und Einfallslot wachsend gleich der Einheit wird. Wir dürfen nach Analogie der Optik oder mit dem Fall der Reflexion am leeren Raum vermuten, daß beim Überschreiten dieses Grenzwinkels in allen Fällen der Energiestrom, den die betreffende Wellenart vorher mit sich geführt hat, sich auf die übrigen noch möglichen Wellen verteilt. Wir können so verfahren, wie es in der Optik meist geschieht, d. h. wir rechnen mit imaginären Größen. Wir erhalten dann imaginäre Amplitudenverhältnisse, die beim Übergang zum Reellen in einer Phasenverschiebung der betreffenden Wellen gegen die einfallende sich äußern. Zugleich finden wir, daß in allen solchen Fällen eine Grenzflächenwelle auftritt, deren Amplitude exponentiell abnimmt mit der Entfernung von der Grenzfläche und die sich längs der Grenzfläche fortpflanzt. Prinzipielle Schwierigkeiten treten in keinem der Fälle auf, solange man sich auf die Betrachtung periodischer Wellen beschränkt. Doch hat es wohl zur Zeit kein erhebliches Interesse, die Formeln für andere Fälle aufzustellen.

Wenn unter hinreichend großen Einfallswinkeln eine Transversalwelle mit Verrückungen in der Einfallsebene auf die Grenzfläche auffällt, so wird der Fall möglich sein, daß sowohl für die reflektierte wie für die durchgehende Longitudinalwelle der Grenzwinkel schon überschritten ist. Wir dürfen alsdann zwei unabhängige Grenzflächenwellen erwarten. Die Energie der einfallenden Transversalwelle aber wird in der reflektierten und in der durchgegangenen Transversalwelle sich wiederfinden. Die Formeln für diesen Fall werden nun freilich schon wieder so umständlich, daß ich mich hier auf den einfacheren beschränke, wo die einfallende Welle eine Transversalwelle mit Verrückungen senkrecht zur Einfallsebene ist. Die Geschwindigkeit $\mathfrak{b}_2$ im Medium 2 sei so groß, daß für Winkel größer als et_0:

$$\frac{\mathfrak{b}_2}{\mathfrak{b}_1} \sin et > 1, \quad \text{also} \quad \cos dt = \sqrt{1 - \left(\frac{\mathfrak{b}_2}{\mathfrak{b}_1} \sin et\right)^2}$$

imaginär wird. Die oben eingeführte Größe k wird

$$k = \frac{\varrho_2 \mathfrak{b}_2 \sqrt{\left(\frac{\mathfrak{b}_2}{\mathfrak{b}_1} \sin et\right)^2 - 1} \, . \sqrt{-1}}{\varrho_1 \mathfrak{b}_1 \qquad \cos et}$$

$$= i \, . \, \varkappa \text{ worin, } \varkappa = \frac{\varrho_2}{\varrho_1} \frac{\mathfrak{b}_2}{\mathfrak{b}_1} \frac{\sqrt{\left(\frac{\mathfrak{b}_2}{\mathfrak{b}_1} \sin et\right)^2 - 1}}{\cos et}$$

$$M_{rt} = \frac{1 - i\varkappa}{1 + i\varkappa} = \frac{1 - \varkappa^2 - 2i\varkappa}{1 + \varkappa^2} = \frac{1 - \varkappa^2}{1 + \varkappa^2} + i \frac{-2\varkappa}{1 + \varkappa^2}$$

$$M_{dt} = \frac{2}{1 + i\varkappa} = \frac{2 - 2\varkappa i}{1 + \varkappa^2} = \frac{2}{1 + \varkappa^2} + i \frac{-2\varkappa}{1 + \varkappa^2}$$

$$\eta_{et} : \mathfrak{R} = 1 \, . \cos \frac{2\pi}{T}\left(t - \frac{\sin et \, x + \cos et \, z}{\mathfrak{b}_1}\right)$$

$$\eta_{rt} : \mathfrak{R} = \frac{1 - \varkappa^2}{1 + \varkappa^2} \cos \frac{2\pi}{T}\left(t - \frac{\sin et \, x - \cos et \, z}{\mathfrak{b}_1}\right)$$

$$+ \frac{+2\varkappa}{1 + \varkappa^2} \sin \frac{2\pi}{T}\left(t - \frac{\sin et \, x - \cos et \, z}{\mathfrak{b}_1}\right)$$

$$\eta_{dt} : \mathfrak{R} = e^{-\frac{2\pi}{T}\sqrt{\left(\frac{\mathfrak{b}_2}{\mathfrak{b}_1} \sin et\right)^2 - 1} \, . \, z} \left\{ \begin{array}{l} \frac{2}{1 + \varkappa^2} \cos \frac{2\pi}{T}\left(t - \frac{\sin et \, x}{\mathfrak{b}_1}\right) \\ \frac{+2\varkappa}{1 + \varkappa^2} \sin \frac{2\pi}{T}\left(t - \frac{\sin et \, x}{\mathfrak{b}_1}\right) \end{array} \right\}$$

$$\eta_{rt} : \mathfrak{R} = 1 . \cos \left\{ \frac{2\pi}{T} \left(t - \frac{\sin et\, x - \cos et\, z}{\mathfrak{b}_1} \right) + \alpha \right\}$$

$$\operatorname{tg} \alpha = \frac{-2\varkappa}{1 - \varkappa^2}$$

$$\eta_{dt} : \mathfrak{R} = e^{-\frac{2\pi}{T} \sqrt{\left(\frac{\mathfrak{b}_2}{\mathfrak{b}_1} \sin et \right)^2 - 1} . z} \frac{2}{\sqrt{1 + \varkappa^2}} \cdot \cos \left\{ \frac{2\pi}{T} \left(t - \frac{\mathfrak{b}_2}{\mathfrak{b}_1} \sin et\, x \right) - \beta \right\}$$

$$\operatorname{tg} \beta = -\varkappa .$$

Diese Formeln zeigen in der Tat, was wir erwarten durften: 1. Die Amplitude der reflektierten Transversalwelle ist gleich der der einfallenden, d. h. die Energie der einfallenden Welle wird total reflektiert. Die reflektierte Welle erfährt eine Phasenänderung um den Winkel α. 2. Die reflektierte Welle zeigt eine Phasenverschiebung von bestimmtem Betrag gegen die einfallende. Statt der durchgehenden Welle haben wir eine nach der x-Richtung fortschreitende, nach der z-Richtung exponentiell abnehmende Grenzflächenwelle.

Magmaschichten.

Nach dem heutigen Stand unserer Kenntnisse über das Erdinnere (vergl. die betreffenden Ausführungen in der gleichzeitig erscheinenden Arbeit „Über Erdbebenwellen“ von E. Wiechert und K. Zoeppritz) müssen wir uns vorstellen, daß die Erde gegenüber schnell wirkenden deformierenden Kräften einen hohen Grad von Steifigkeit besitzt. Dies lehrt uns vor allem die Verfolgung der Scherungswellen durch das Erdinnere hindurch. Trotz dieses im allgemeinen sicher stehenden Resultates wäre es wohl denkbar, daß eine Schicht verminderter Steifigkeit oder gar eine zähflüssige Magmaschicht in geringer Tiefe vorhanden wäre. Eine solche würde sich nämlich im Verlauf der Laufzeitkurve wenig bemerkbar machen. Und aus diesem Verlauf war ja gerade das oben angegebene Resultat bezüglich der Steifigkeit des Erdinnern hergeleitet worden. Sind wir nun vorläufig außer Stande, auf Grund der Laufzeitkurve etwas Sicheres in dieser Frage zu erfahren, so lohnt es sich doch wohl zu prüfen, in welcher Weise denn Transversalwellen eine solche, eben charakterisierte Schicht zu durchdringen vermögen. Weniger geeignet, eine Entscheidung in der vorliegenden Frage herbeizuführen, sind Transversalwellen mit Verrückungen in der Einfallsebene. Denn solche Wellen

werden immer auch dann erregt, wenn eine Longitudinalwelle eine Unstetigkeitsfläche durchsetzt. Viel klarer dagegen liegen die Verhältnisse bei Transversalwellen mit Schwingungen senkrecht zur Einfallsebene. Sie können nicht durch Longitudinalwellen ausgelöst werden und müssen deshalb auf ihrem Wege durch das Erdinnere auch immer gewissermaßen ihre Selbständigkeit bewahren. Nun ist schon wiederholt darauf hingewiesen worden, daß in den „zweiten Vorläufern“ die Schwingungen senkrecht zur Einfallsebene besonders stark hervortreten. Wir haben uns also die Frage zu beantworten: wie vermögen Transversalwellen mit Schwingungen senkrecht zur Einfallsebene eine wenn auch dünne Magmaschicht zu durchdringen? Weiter muß sich dann die Frage daran knüpfen, was wir über das Vorhandensein einer solchen Magmaschicht aus dem Umstand schließen können, daß Transversalwellen der gedachten Art sie durchdringen.

Ich berücksichtige zunächst den Fall einer flüssigen, wenn auch sehr zähflüssigen Schicht und werde dann noch kurz auf den Fall einer elastischen, aber durch geringere Steifigkeit ausgezeichneten Schicht eingehen. Es sind hier die Beziehungen zwischen den Gleichungen der Elastizitätstheorie und denen der Hydrodynamik von Wichtigkeit. Denken wir uns die Magmaschicht als kompressible Flüssigkeit, jedoch als sog. ideale Flüssigkeit, d. h. ohne innere Reibung, so sind Transversalwellen überhaupt nicht im Stande, sie zu durchdringen. Diese Möglichkeit haben wir also von vornherein auszuschließen. Anders liegt die Sache bei reibenden Flüssigkeiten. Hier sind in der Tat Transversalwellen möglich. Wir können direkt eine Lösung, die von Stokes herrührt[1]), auf unseren Fall ebener Wellen anwenden. Die hydrodynamischen Gleichungen mit Berücksichtigung der Reibung reduzieren sich im Falle ebener Transversalwellen auf die aus der linearen Wärmeleitung bekannte Gleichung:

$$\frac{\partial v}{\partial t} = \frac{\mu}{\varrho} \frac{\partial^2 v}{\partial x^2},$$

worin μ der Koeffizient der inneren Reibung und ϱ die Dichte ist.

Ist v, die transversale Verrückung, proportional mit $e^{\frac{2i\pi t}{T}}$, so nimmt die Lösung die Form an:

$$v = M . e^{-\gamma x} . \cos \frac{2\pi}{T}\left(t - \gamma \frac{T}{2\pi} x\right), \text{ worin } \gamma = \sqrt{\frac{\pi . \varrho}{T . \mu}}.$$

1) Vergl. Rayleigh, Theorie des Schalls. Erste Aufl. § 347.

Die Transversalwelle pflanzt sich also, wenn auch mit exponentiell abnehmender Amplitude in die Flüssigkeit hinein fort. Die Fortpflanzungsgeschwindigkeit beträgt:

$$\mathfrak{b}_1 = \frac{2\pi}{T}\frac{1}{\gamma} = \sqrt{\frac{4\pi \, . \, \mu}{T \, . \, \varrho}},$$

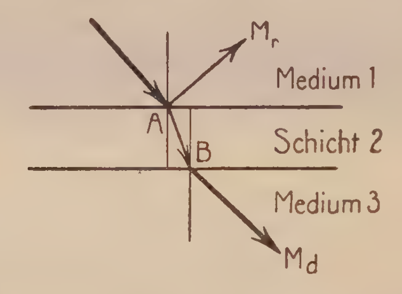
wächst also mit der Quadratwurzel aus dem Reibungskoeffizienten und mit derjenigen aus der Frequenz $n = \frac{2\pi}{T}$. Auch die Dämpfung auf die gleiche Wegstrecke x nimmt zu mit wachsendem Reibungskoeffizienten. Für eine Flüssigkeit von der Zähigkeit des Wassers beträgt die Fortpflanzungsgeschwindigkeit für $T = 10$ Sekunden ca. 1 mm, für Glyzerin einige Zentimeter. Dabei sinkt die Amplitude in Wasser schon bei 1 cm Weglänge auf etwa $^1/_{400}$ ihres ursprünglichen Wertes. In Glyzerin erfolgt die gleiche Abnahme auf etwa 60 cm Weglänge. Wir müssen also eine ganz andere Größenordnung für die Zähigkeit der Magmaschicht zu Grunde legen, falls überhaupt Transversalwellen durch sie hindurch gehen sollen.

Den Reibungskoeffizienten von Pech bei 0° hat G. H. Darwin[1]) der Größenordnung nach zu 10^{11} bestimmt. Damit erhält man eine Fortpflanzungsgeschwindigkeit von nahezu 3 km/sec. für eine Periode von $T = 10$ Sekunden. Die Amplitude sinkt dabei in etwa 4 km Wegstrecke auf $\frac{1}{e} = 0.37$. Wie man sieht, nähert sich jetzt die Fortpflanzungsgeschwindigkeit bei mittleren seismischen Perioden der Geschwindigkeit elastischer Transversalwellen in festem Gestein (ca. 4 km/sec.). Es dringt nun ein erheblicher Bruchteil der Welle in die Magmaschicht ein, während bei dünnflüssigem Magma auch abgesehen von der Dämpfung schon wegen des starken Sprunges in der Geschwindigkeit nahezu die ganze Wellenenergie beim Auftreffen auf die Grenzfläche reflektiert werden müßte. Diese Schwierigkeit fällt somit bei einem äußerst zähen Medium fort. Wir haben demnach als Unterschiede gegenüber elastischer Wellenbewegung nur 1. die Abhängigkeit der Fortpflanzungsgeschwindigkeit von der Schwingungsperiode und 2. die starke Dämpfung durch innere Reibung. Ganz besonders muß uns dabei die enorme Dämpfung selbst bei großer Zähigkeit auffallen. Und wir erhalten vorläufig das Resultat: Wenn eine

1) Phil. Transakt. **170**, S. 1, 1878.

zähflüssige Magmaschicht in gewisser Tiefe vorhanden ist, so kann sie doch jedenfalls nur sehr geringe Mächtigkeit (der Größenordnung nach vielleicht 10 km) haben, widrigenfalls Transversalwellen sie nicht in merklicher Stärke zu durchdringen vermögen.

Versuchen wir nun in zweiter Linie eine Abschätzung der Durchlässigkeit für Transversalwellen unter Zugrundelegung einer festen elastischen Schicht, jedoch mit verminderter Steifigkeit. In diesem Fall wird möglicherweise der Einfluß der inneren Reibung viel geringer sein. Wir vernachlässigen ihn deshalb ganz und betrachten die Schwächung einer Transversalwelle beim Durchgang durch eine weniger elastische Schicht allein auf Grund der Reflexion an der Grenzfläche. Ich denke mir also zwischen Gestein von normalen elastischen Eigenschaften eine Schicht von verminderter Steifigkeit eingeschaltet. Was die Materialkonstante anlangt, so



sind die Dichteunterschiede jedenfalls nur sehr gering. Ich setze deshalb $\varrho_1 = \varrho_2 = \varrho_3$. Für die Berechnung der Amplitudenverhältnisse haben wir die oben für den Fall einer einfallenden Welle mit Verrückungen senkrecht zur Einfallsebene aufgestellten Formeln anzuwenden. Die Stärke der bei A ins Medium 1 rückwärts reflektierten und der bei B ins Medium 3 eindringenden Welle hängt bei konstanter Dichte in allen drei Medien allein von der Geschwindigkeitsänderung beim Übergang von 1 nach 2 und von 2 nach 3 ab. Ich messe deshalb die Steifigkeitsverminderung in der Schicht 2 direkt durch die Geschwindigkeitsänderung, wobei ich noch $\mathfrak{b}_1 = \mathfrak{b}_3$ setze. Die folgende kleine Tabelle gibt für einige Einfallswinkel die gewünschten Amplituden M_r und M_d und zwar in $^0/_0$ der Amplitude der bei A einfallenden Welle. Die Amplituden sind berechnet für die in der ersten Zeile angegebene Geschwindigkeitsabnahme. Die Tabelle gestattet auch unmittelbar die Anwendung auf Reibungswellen gleicher Geschwindigkeit in einer zähen Flüssigkeit. Die Reibungskoeffizienten, die bei einer

Periode von 10 Sek. den betreffenden Geschwindigkeitswerten sich zuordnen, sind in der nächsten Zeile angegeben. Dabei ist die Dichte = 2,7 gesetzt. Schließlich ist noch die Wegstrecke in einer weiteren Zeile angegeben, auf die durch Dämpfung in einer Flüssigkeitsschicht der angegebenen Zähigkeit die Amplitude der Transversalwelle auf $1/e = 0{,}37$ heruntersinkt:

Tabelle.

Einfallswinkel	$\mathfrak{b}_1 : \mathfrak{b}_2 = 4:4$		$\mathfrak{b}_1 : \mathfrak{b}_2 = 4:3$		$\mathfrak{b}_1 : \mathfrak{b}_2 = 4:1$		$\mathfrak{b}_1 : \mathfrak{b}_2 = 4:0{,}4$		$\mathfrak{b}_1 : \mathfrak{b}_2 = 4:0{,}04$		$\mathfrak{b}_1 : \mathfrak{b}_2 = 4:0$	
	M_r	M_d	M_r	M_d	M_r	M_d	M_r	M_d	M_r	M_d	M_r	M_d
90°	−1	0	−1	0	−1	0	−1	0	−1	0	−1	0
60°	0	1	−0,06	0,99	0,34	0,88	0,60	0,64	0,95	0,098	1	0
30°	0	1	+0,11	0,99	0,55	0,69	0,75	0,44	0,97	0,057	1	0
0°	0	1	+0,14	0,98	0.60	0.65	0,82	0,33	0,98	0,039	1	0
Reibungskoeffizient	$3{,}4 \cdot 10^{11}$		$1{,}9 \cdot 10^{11}$		$2{,}1 \cdot 10^{10}$		$2{,}1 \cdot 10^{8}$		$2{,}1 \cdot 10^{6}$		0	
Dämpfungsstrecke	6,3 km		4,7 km		1,6 km		0,16 km		0,016 km		0	

Die Tabelle lehrt uns, eine wie starke Schwächung eine Transversalwelle beim Durchgang durch eine Schicht verminderter Steifigkeit oder durch eine entsprechende, zähflüssige Schicht erfährt. Im letzteren Fall tritt außerdem noch die Schwächung durch innere Reibung hinzu, die so stark ist, daß sie bei nicht ganz besonders hohem Grad von Zähigkeit jeden Durchgang einer Transversalwelle unmöglich machen muß. Nicht vergessen dürfen wir ferner, daß eine zu einer fernen Station gelangende transversale Welle die fragliche Schicht zweimal, nämlich beim Eindringen und bei der Rückkehr aus dem Erdinnern, zu passieren hat, daß also die Schwächung sich für die zur Erdoberfläche heraufdringende Welle noch mit sich selbst multipliziert.

Halten wir hiermit allein das Resultat zusammen, daß in den „2ten Vorläufern" die Transversalwellen quer zur Einfallsebene meist sehr kräftig hervortreten, so erscheint es wenig wahrscheinlich, daß eine solche Schicht überhaupt in allgemeiner Verbreitung vorhanden ist. E. Wiechert hat kürzlich[1]) zur Erklärung transversaler Oberflächenwellen auf eine solche Schicht größerer Nach-

1) E. Wiechert und K. Zoeppritz, l. c.

giebigkeit aufmerksam gemacht. Darnach wäre eine äußere feste Gesteinschicht von etwa 25—35 km. Dicke wenigstens am Herd solcher Erdbeben, die von langperiodigen transversalen Oberflächenwellen begleitet sind, vom starren Erdkern durch eine nachgiebige Magmaschicht getrennt und daher im Stande, in freie Querschwingungen zu geraten wie eine beiderseits durch „freie" Ebenen begrenzte Platte. Auf Grund unsrer Formeln können wir diese Hypothese nun daraufhin prüfen, ob die Grenze der äußeren festen Kruste gegen die vermutete Magmaschicht wirklich als „frei" gelten kann, wenn doch gleichzeitig Transversalwellen mit Verrückungen senkrecht zur Einfallsebene durch die Schicht hindurchgehen sollen. Freie Querschwingungen einer Platte sind offenbar nur dann möglich, wenn die Schwingung von beiden Begrenzungsebenen der Platte fast ohne Schwächung in die Platte hinein zurückreflektiert wird. Während jeder Schwingung fällt somit gewissermaßen eine Transversalwelle unter dem Einfallswinkel 0 auf die untere Grenzfläche auf. Wird hier nur ein Teil der Energie reflektiert, so muß der Schwingungsvorgang in der Platte je nach dem Anteil, der in die unterliegende Magmaschicht eingedrungen ist, mehr oder weniger stark gedämpft verlaufen. Ziehen wir nun unsere Tabelle auf S. 81 heran, so sehen wir, daß auch bei einer Geschwindigkeitsabnahme in der tieferliegenden Schicht auf den zehnten Teil (von 4 km/sec. auf 0,4 km/sec.) die Amplitude der reflektierten Welle doch um ca. 1/5 bei jedesmaliger Reflexion abnimmt. Also nach 3 Schwingungen oder bei einer Periode von 18 Sekunden noch vor Ablauf einer Minute wäre die Amplitude der „freien" Querschwingung auf etwa die Hälfte gesunken. Um die Entstehung einer längeren Folge von transversalen Oberflächenwellen mit ganz allmählich sinkender Amplitude auf Grund freier Plattenschwingungen erklären zu können, müßten wir also eine noch viel stärkere Geschwindigkeitsabnahme von der festen äußeren zur nachgiebigen inneren Schicht annehmen, etwa von 4 auf 0,04 km/sec. In diesem Fall würde aber, wie unsere Tabelle lehrt, die Amplitude der durchgelassenen Transversalwelle bei einmaligem Passieren der Schicht auf etwa 4%, bei zweimaligem sogar auf noch nicht 2‰ heruntersinken. Die Transversalwelle mit Schwingungen senkrecht zur Einfallsebene würde sich dann, weil zu schwach, der Registrierung wohl ganz entziehen. Mir scheint deshalb das Auftreten kräftiger Schwingungen senkrecht zur Einfallsebene in den „2ten Vorläufern" nicht verträglich mit der Annahme einer nachgiebigen Schicht in geringer Tiefe unterhalb der Herdregion. Denken wir uns die geringere Steifigkeit der

Schicht verursacht durch ein Erweichen der Gesteine, so wird außer der Geschwindigkeitsabnahme für Scherungswellen auch noch innere Reibung sich geltend machen, die das Bild noch ungünstiger gestaltet, denn sie bedingt gleichzeitig eine weitere Schwächung der durchgelassenen Welle und eine verstärkte Dämpfung der Oberflächenschichtwelle.

Unser Resultat wäre also: Keine allgemein verbreitete Magmaschicht; es erscheint nicht unwichtig, darauf hinzuweisen, daß dieses Ergebnis in guter Übereinstimmung ist mit dem, was die Geologie uns lehrt. Denn auch sie betrachtet neuerdings aus mehreren Gründen die flüssigen Magmaherde, deren sichtbare Spuren uns in den Vulkanen entgegentreten, als lokalisierte Phänomene innerhalb der festen Erdrinde und ohne Zusammenhang mit einem etwa allgemein verbreiteten Magma. Von großem Interesse müßte es sein, in möglichster Nähe eines großen, noch tätigen Vulkans messende Erdbebenbeobachtungen anzustellen. Hier wäre möglicherweise der dämpfende Einfluß auf die Vorläuferwellen zu erkennen, den das Magma in der Tiefe, falls der Flüssigkeitsherd genügende Ausdehnung hat, ausüben muß.

Ganz offen bleibt bei unserer Auffassung die Frage nach der Entstehung transversaler Oberflächenwellen. Doch wird man sagen müssen, daß das Beobachtungsmaterial in dieser Richtung noch einer weiteren vergleichenden Bearbeitung bedarf. Es wäre in dieser Hinsicht außerordentlich erwünscht, daß von stärkeren Erdbeben die Aufzeichnungen aller hinreichend vergrößernden Seismographen, die zur Messung der wahren Bodenbewegung geeignet sind, also vorzüglich der gedämpften, oder solcher ungedämpften Apparate, bei denen die Eigenperiode alle Bodenperiode übertrifft, durch die Zentralstellen der Erdbebenorganisation vervielfältigt herausgegeben würden. Damit Hand in Hand muß natürlich eine zweckmäßige Erweiterung des Beobachtungsnetzes gehen.

Bezüglich der Deutung der transversalen Oberflächenwellen sei noch daran erinnert, daß die Voraussetzung, unter der das Huygens'sche Prinzip, auf Erdbebenstrahlen angewandt, einen geradlinigen Bebenstrahl liefert, nämlich die Voraussetzung gleicher Geschwindigkeit in allen benachbarten Punkten, für die Oberflächenwellen nicht sehr weitgehend erfüllt sein wird. Man darf wohl annehmen, daß Unterschiede der elastischen Eigenschaften des Erdmaterials, besonders in den äußeren Schichten der Oberfläche, sich geltend machen. Dieser Umstand, sowie Unregelmäßigkeiten des Oberflächenreliefs können dazu führen, daß vermutlich mehr als bei den Vorläuferwellen gerade bei den Oberflächenwellen die

Schwingungsrichtung, auch wenn die Wellenart longitudinalen Charakter hat, stark und mit der fortschreitenden Ausbreitung der Wellen auch wechselnd von der Richtung Herd—Erdbebenstation abweicht. Im ganzen haben wir es hier mit Problemen zu tun, deren endgültige Erledigung erst die Zukunft bringen wird.

Göttingen, Geophysikal. Institut. 22. Dez. 1907.

Die Elastizitätskonstanten von kristallisiertem Kalialaun.

Von

W. Voigt.

Vorgelegt in der Sitzung vom 21. Februar 1919.

1. Die nachstehend mitgeteilten und berechneten Beobachtungen an Kalialaun sind vor mehr als 30 Jahren ausgeführt worden, etwa gleichzeitig mit denjenigen an Flußspat, Pyrit u. s. f.[1]); ihre Mitteilung ist damals unterblieben, weil einige Zahlen (besonders für die Biegungsmoduln) derart aus dem Rahmen der übrigen herausfielen, daß die Gesamtheit der Resultate mir dadurch unbrauchbar erschien. Jetzt, wo alle Zahlenwerte von Elastizitätskonstanten an Interesse gewonnen haben, nahm ich die alten Messungsprotokolle, wie auch die erhaltenen Präparate noch einmal vor und entdeckte bei genauerer Prüfung, daß nicht, wie ich vermutete, gestörtes Kristallmaterial, sondern ein einfaches Versehen bei einigen Messungen der Stäbchen*breiten* (Fehler um ganze Umdrehungen der Mikrometerschraube) die wichtigsten Unstimmigkeiten veranlaßt hatte, und daß nach Berichtigung dieser Fehler ein System von Werten entsteht, das die elastischen Parameter von Alaun so genau zu liefern vermag, als bei der etwas heikeln Substanz nur irgend zu hoffen war.

Die Hauptschwierigkeit einer genaueren Bestimmung der Elastizitätskonstanten von so überaus weichen Materialien wie Alaun erwächst durch den Umstand, daß Arbeiter, die vorwiegend härteres Material (z. B. Glas) bearbeiten, mit der Herstellung von regelmäßig gestalteten Präparaten aus weichem Material meistens

---

1) W. Voigt, Gött. Nachr. 1888 No. 11.

nicht recht zustande kommen. So weisen auch die zur Beobachtung gelangten Alaunstäbchen (von Dr. Steeg u. Reuter hergestellt) ziemlich stark gekrümmte Begrenzungsflächen auf, — die u. A. die Anwendung der neuen Methoden[1]) der Newtonschen Ringe von vorn herein ausschließen und die Anwendung der älteren, s. Z. beschriebenen[2]) erfordern.

Um zu befriedigenden Resultaten zu kommen muß man in solchen Fällen insbesondere die Dicken der Stäbchen, die in die Biegungs- und Drillungsformeln zur dritten Potenz eingehen, an sehr vielen Stellen beobachten und ihre Unregelmäßigkeit bei den Biegungsmessungen wirklich theoretisch in Rechnung setzen. Denn bei den Biegungen des in der Mitte belasteten Stabes geben die verschiedenen Teile längs der Achse einen verschiedenen Anteil zur Senkung der belasteten Mitte, während bei den Drillungen alle Teile zu der Drehung des Endquerschnittes gleichmäßig beitragen. Von den Breiten, die lineär in die Formeln eingehen und überdies relativ schwächer variieren, genügt es in beiden Fällen, den arithmetischen Mittelwert der an einer größeren Anzahl von Stellen ausgeführten Messungen für das benutzte Bereich der Länge in Rechnung zu setzen.

Die Messungen der Dicken geschahen mit Hilfe eines früher beschriebenen Schraubenmikrometers von rund 0,5 mm Ganghöhe, 500 Teilen der Trommel und genauem Wert des Trommelteiles = 1/992,6 mm, u. zw. in sieben oder (bei einigen kürzeren Stäbchen) fünf äquidistanten Querschnitten. In jedem Querschnitte wurde an zwei der Mitte und zwei den Rändern nahen Stellen gemessen, so daß 28 resp. 20 Zahlen gewonnen wurden. Die auf jeden Querschnitt bezüglichen wurden durch Mittelbildung vereinigt. Die Breiten wurden in denselben Querschnitten in zwei bez. der Längsachse um 180° gedrehten Lagen im Mikrometer gemessen.

2. Die Theorie der Berücksichtigung der veränderlichen Dicke bei den Biegungsmessungen ist sehr einfach. Es ist bekannt, daß man die beiden Hälften eines an den Enden aufliegenden und in der Mitte belasteten Stabes als zwei Einzelstäbe zu behandeln hat. Wir setzen demgemäß, indem wir die z-Koordinate von der Mitte des Stabes nach der einen wie nach der andern Seite hin positiv rechnen, für die Dicke D das Gesetz an

$$D = D_0(1 \pm az + bz^2), \quad (1)$$

---

1) W. Voigt, Gött. Nachr. 1918 No. 3.

2) W. Voigt, l. c. 1887 No. 19.

wobei das $\pm$ sich auf die beiden Seiten bezieht, und der Klammerausdruck sich nur wenig von Eins unterscheidet.

Die Differentialgleichung der Biegungskurve schreiben wir

$$2)\qquad \frac{BD^3}{12}\frac{d^2u}{dz^2} = s'_{33}(l-z)\Pi,$$

wobei l die Länge der Stabhälfte, Π die von der Unterstützung bei $z = l$ ausgeübte Reaktionskraft bezeichnet, und s'_{33} der Dehnungsmodul parallel der Stabachse ist.

Durch Entwicklung des Ansatzes (1) für D gewinnen wir aus (2)

$$3)\qquad \frac{BD_0^3}{12}\frac{d^2u}{dz^2} = s'_{33}(l-z)\Pi(1 \mp 3az - 3cz^2),$$

wobei $c = b - 2a^2$. Diese Gleichung ist zu integrieren unter der Annahme, daß du/dz bei $z = 0$ für die beiden Stabhälften entgegengesetzt gleiche Werte besitzt, und u bei $z = 0$ verschwindet. Die Ausweichung η der Enden gegenüber der Mitte (oder umgekehrt) findet sich dann gegeben durch

$$4)\qquad \frac{\eta BD_0^3}{4} = s'_{33}\Pi l^3\left(1 - \frac{3cl^2}{10}\right),$$

oder, wenn man

$$D_0\left(1 + \frac{cl^2}{10}\right) = \overline{D},\ \Pi = \tfrac{1}{2}P,\ l = \frac{L}{2}$$

setzt, durch

$$5)\qquad \eta = \frac{s'_{33}L^3P}{4B\overline{D}^3}.$$

Hierin ist P die Belastung der Stabmitte, L die ganze Stablänge.

Beobachtet man die mittlere Stabdicke in Abständen $z = hl/n$, wobei h und n ganze Zahlen sind, so kann man statt (1) schreiben

$$6)\qquad D = D_0 + h\delta_1 + h^2\delta_2$$

und erhält, da die Vergleichung

$$\delta_1 = \frac{alD_0}{n},\ \delta_2 = \frac{blD_0}{n}$$

ergibt,

$$7)\qquad \overline{D} = D_0 + \frac{n^2}{10}\left(\delta_2 - \frac{2\delta_1^2}{D_0}\right).$$

Dies ist der Ausdruck der nach (5) für die Berechnung der Biegungsbeobachtungen einzuführenden mittleren Stäbchendicke.

Die Messung der durch (5) bestimmten Durchbiegung des Stäbchens geschah in der früher beschriebenen Weise mit Fernrohr, Spiegel und Skala. Ein Skalenteil entsprach $2{,}954.10^{-4}$ mm.

Das Material für die Herstellung der Alaunstäbchen lieferten mehrere große oktaedrische Kristalle, die, wie gewöhnlich, starke Schichtung nach Oktaederflächen zeigten und in merklich ungestörter Substanz außer den Bereichen um die Kristallkanten nur zentrale den Würfelebenen parallele Lamellen darboten. Diese letzteren Bereiche wurden zur Herstellung einiger Stäbchen benutzt, deren Länge, Breite und Dicke zu Würfelachsen parallel waren; diese Präparate sind mit W bezeichnet. Besseres Material boten die Bereiche der Oktaederkanten dar. Hier gelang es, eine ziemliche Anzahl von Stäbchen herzustellen, deren Längsachse zu der Oktaederkante parallel, also zu einer Granatoederfläche normal lag, während die Breite normal, die Dicke parallel einer Oktaederfläche war. Diese Gattung ist mit G bezeichnet.

Werden die Hauptmoduln für das reguläre System durch s_{11}, s_{23}, s_{44} bezeichnet, so sind die Dehnungs- oder Biegungsmoduln für diese beiden Gattungen Präparate resp.

$$8)\qquad (s'_{33})_w = s'_w = s_{11},\ (s'_{33})_g = s'_g = \tfrac{1}{4}\left(2(s_{11}+s_{23})+s_{44}\right).$$

3. Der Drillung sind, der bessern Qualität und der größeren Anzahl wegen, wesentlich nur die Stäbchen von der Gattung G unterzogen worden, obgleich für die beschriebene Orientierung die Theorie der bez. Deformation nicht in Strenge entwickelt werden kann. Immerhin läßt sich, wie ich an anderer Stelle[1]) ausgeführt habe, leicht eine Annäherungsformel von hinreichender Genauigkeit gewinnen. Fällt nämlich, wie hier, die Stabachse in eine zweizählige kristallographische Symmetrieachse, und ist für ein System $X'\ Y'\ Z'$ parallel $D\ B\ L$ der Modul s'_{45} klein neben $\sqrt{s'_{44}.s'_{55}}$, dann nimmt, wenn $\frac{1}{2}(s'^2_{45}/s'_{44}s'_{55})^2$ in dem zweiten Glied des Nenners neben Eins vernachlässigt wird, die Formel für den Drehungswinkel ϑ des Querschnittes $z = L$ gegen den Querschnitt $z = 0$ die Gestalt an

$$9)\qquad \vartheta = \frac{3\,LNs'_{44}}{BD^3\left(1-0{,}630\,\frac{D}{B}\sqrt{\frac{s'_{44}}{s'_{55}}}\right)}.$$

Dabei ist, wenn die L- resp. Z'-Richtung in eine Granatoedernormale, die B- resp. X'-Richtung in eine Oktaedernormale fällt, das System der Richtungskosinus gegen die XYZ-Hauptachsen gegeben durch

1) W. Voigt, Kristallphysik p. 649 u. f.

10)

	x'	y'	z'
x	$\frac{1}{\sqrt{3}}$	$-\sqrt{\frac{2}{3}}$	0
y	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{6}}$	$\frac{1}{\sqrt{2}}$
z	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{6}}$	$\frac{1}{\sqrt{2}}$,

und für die Moduln s'_{44}, s'_{55}, s'_{45} gelten die Ausdrücke[1])

11) $$s'_{44} = \tfrac{2}{3}(s_{11} - s_{23} + s_{44}),\ s'_{55} = \tfrac{1}{3}(4(s_{11} - s_{23}) + s_{44}),$$
$$s'_{45} = (2(s_{11} - s_{23}) - s_{44})\,\tfrac{1}{3}\sqrt{2}.$$

Für isotrope Körper ist $2(s_{11} - s_{12}) = s_{44}$ also $s'_{44} = s'_{55} = s_{44}$ $s'_{45} = 0$; je geringer die elastische Anisotropie eines regulären Kristalles ist, um so näher werden diese letzteren Beziehungen auch bei ihm erfüllt sein. Der relativ geringe Unterschied zwischen s'_g und s'_w, den, wie unten zu zeigen, die Biegungsbeobachtungen ergeben haben, läßt erkennen, daß ein solches Verhältnis bei Alaun vorliegt.

Die Beobachtung der durch (9) bestimmten Drehung ϑ geschah mit Hülfe von zwei in den Querschnitten $z = 0$ und $z = L$ auf dem Stäbchen befestigten Spiegeln durch Fernrohr und Millimeterskala. Bezeichnet σ die Anzahl mm, um die sich die Bilder in beiden Spiegeln gegen einander verschieben, E den Abstand Spiegel-Skala, so ist

12) $$\vartheta = \sigma/2E.$$

Die Millimeterskala war um 0,00374 zu groß; E betrug 5173 mm.

Das Drehungsmoment N drückt sich durch Belastung P und Hebelarm H aus gemäß der Formel $N = HP$; für H war der Wert 36,8 mm zu benutzen.

Die Drillungsbeobachtungen wurden zur Vergrößerung der Genauigkeit bei Festhaltung einmal des linken, einmal des rechten Stabendes, also Drehung am rechten resp. linken Ende ausgeführt, wie das in den folgenden Tabellen durch die Buchstaben rR und lR ausgedrückt ist. In jeder Stellung wurden mehrere Belastungen angewandt, z. B. das Gewicht der Wagschale S und 10 resp. 20 gr. Zur Elimination der Achsenreibung wurden die Gleichgewichtslagen von der Seite größerer und kleinerer Belastung erreicht.

1) Die Ableitung dieser und analoger Ausdrücke ist im Anhang auseinandergesetzt.

Die so erhaltenen Zahlen σ gestatteten dann durch Kombination den wirklichen Wert σ_{10} oder σ_{20}, der 10 oder 20 gr Belastung entspricht, abzuleiten, obwohl die Beobachtung der Einstellung ohne jede Belastung durch Reibungswirkung gefälscht bleibt.

4. In der folgenden Tabelle der Dimensionen sind für das erste Präparat zur Orientierung über die Veränderlichkeit der Dimensionen längs der Achse und längs der einzelnen Querschnitte des Stäbchens sämtliche Messungen angegeben, weiterhin nur die Mittel für die Querschnitte, dazu gemäß den Formeln (6) und (7) die aus den Dicken-Mitteln berechneten δ_1 und δ_2. Einheiten sind die Trommelteile des Mikrometers (jedes = 1/992,6 mm).

I. Dimensionen.

G. No. 1. $D = 1000 + \delta$.

δ =	164,1	149,2	135,4	121,9	109,6	101,5	90,3
	164,8	151,0	135,7	124,0	112,4	104,7	92,2
	167,1	150,5	135,4	123,3	113,2	104,5	92,2
	169,7	154,9	139,5	126,6	114,5	104,5	92,8
Mittel:	166,4	151,4	136,0	124,0	112,6	103,8	91,9
ber.:	166,2	150,8	136,7	124,0	112,5	102,4	93,6

$\delta_0 = 124{,}0 \quad \delta_1 = 12{,}1 \quad \delta_2 = 0{,}65.$

$B = 5900 + \beta$.

β =	44,6	44,8	44,8	44,7	44,3	43,5	42,1
	44,2	44,8	44,9	44,7	44,2	43,5	41,8
Mittel:	44,4	44,8	44,8	44,7	44,2	43,5	42,0.

Man sieht, daß der Ansatz (1) resp. (6), nach dem die unter ber. angegebenen Zahlen berechnet sind, den wirklichen Verlauf der Dicke sehr gut wiedergiebt, und darf daraus schließen, daß die Berücksichtigung der Veränderlichkeit der Dicke durch die S. 87 entwickelte Theorie eine befriedigende Genauigkeit liefern wird.

G. No. 2. $D = 1000 + \delta$

δ =	170,4	159,7	141,7	124,8	104,9	88,9	64,8
ber.:	172,4	157,8	141,8	124,6	106,2	86,6	65,6

$\delta_0 = 124{,}6 \quad \delta_1 = 17{,}8 \quad \delta_2 = -0{,}62.$

$B = 5900 + \beta$.

$\beta = 52 \quad 84 \quad 73 \quad 63 \quad 58 \quad 48 \quad 29$

G. No. 3. $D = 1000 + \delta$.

δ =	177,4	165,1	146,7	131,8	117,1	105,1	86,7
ber.:	179,0	162,9	147,2	132,0	117,2	102,7	88,8

$\delta_0 = 132{,}0 \quad \delta_1 = 15{,}0 \quad \delta_2 = 0{,}21.$

$B = 5900 + \beta$.

$\beta = 33 \quad 51 \quad 61 \quad 65 \quad 69 \quad 69 \quad 54.$

G. No. 4. $D = 1000 + \delta$.

δ = 173,4	164,8	151,4	136,6	121,0	104,3	84,5
ber.: 174,7	163,7	151,1	136,9	121,1	103,7	84,7

$\delta_0 = 136{,}9 \quad \delta_1 = 15{,}0 \quad \delta_2 = -0{,}80.$

$B = 5900 + \beta$.

β = 32 48 59 64 66 65 56

G. No. 5. $D = 1000 + \delta$.

δ = 164,9	157,9	145,3	133,3	120,5	108,4	98,2
ber.: 167,1	156,1	144,8	133,2	121,4	109,3	96,9

$\delta_0 = 133{,}3 \quad \delta_1 = 11{,}7 \quad \delta_2 = -0{,}13.$

$B = 5900 + \beta$

β = 58 76 78 74 72 72 50.

G. No. 6. $D = 1000 + \delta$.

δ = 159,7	148,3	137,5	126,1	114,2	100,0	82,7
ber.: 158,5	148,9	138,1	126,3	113,3	99,3	84,1

$\delta_0 = 126{,}3 \quad \delta_1 = 12{,}4 \quad \delta_2 = -0{,}55.$

$B = 5900 + \beta$.

β = 27 51 50 54 57 61 58.

W. No. 1. $D = 1100 + \delta$.

δ = 172,0	163,2	146,2	135,0	115,0	101,0	82,4
ber.: 173,5	161,1	147,5	132,9	117,1	100,3	82,3

$\delta_0 = 132{,}9 \quad \delta_1 = 15{,}2 \quad \delta_2 = -0{,}56.$

$B = 5700 + \beta$.

β = 57 75 83 81 76 60 15.

W. No. 2. $D = 1200 + \delta$.

δ = 123,6	116,4	110,6	86,5	69,4	53,2	34,4
ber.: 125,7	114,0	110,9	86,3	70,3	52,8	33,9

$\delta_0 = 86{,}3 \quad \delta_1 = 15{,}3 \quad \delta_2 = -0{,}72.$

$B = 5500 + \beta$.

β = 13 25 26 29 27 16 −3.

W. No. 3. $D = 700 + \delta$.

δ = 36,5	31,7	29,7	29,3	26,3	27,4	29,3
ber.: 36,5	32,5	29,5	27,7	27,1	27,5	29,1

$\delta_0 = 27{,}7 \quad \delta_1 = 1{,}24 \quad \delta_2 = 0{,}57.$

$B = 6000 + \beta$.

β = 113 123 120 121 114 94 65.

W. No. 4. $D = 900 + \delta$.

δ = 13,7	15,4	14,4	13,5	13,8	13,5	21,2
ber.: 18,4	15,9	14,3	13,5	13,5	14,3	15,8

$\delta_0 = 13{,}5 \quad \delta_1 = 0{,}42 \quad \delta_2 = 0{,}40.$

$B = 6000 + \beta$.

β = 88 82 75 52 35 15 −6.

W. No. 5. $D = 700 + \delta.$

$\delta = 37{,}7$	37,7	40,6	42,8	40,8
ber.: 36,4	38,9	40,6	41,7	42,0

$\delta_0 + 40{,}6 \quad \delta_1 = 1{,}4 \quad \delta_2 = -0{,}34.$

$B = 6000 + \beta.$

$\beta = 96 \quad 116 \quad 124 \quad 128 \quad 131.$

5. Die nachstehende Zusammenstellung der Resultate der Biegungsbeobachtungen bringt in der ersten Zeile nach der Ordnungsnummer des Stäbchens dessen für die Berechnung ev. nach S. 87 reduzierte Dimensionen, Dicke, Breite, Länge, sowie die an der Mitte des Stäbchens angebrachte Belastung. Die zweite Zeile enthält unter η_1 und η_2 die dieser Belastung entsprechenden Durchbiegungen der Stäbchenmitte in zwei gegeneinander verwendeten Lagen; ferner unter η' die Korrektion, die an diesen Zahlen wegen der Eindrückung an den Unterstützungsstellen anzubringen ist. Diese η' sind durch besondere Beobachtungen ermittelt, bei denen einige Stäbchen bei der sehr geringen Länge L' von 14 mm mit demselben Gewicht P gebogen wurden; die Kombination dieser Messungen mit den bei der Länge L ausgeführten gestattet (da die wirklichen Biegungen mit L'^3 und L^3 proportional sind und die Eindrückungen bei gleicher Belastung mit Wahrscheinlichkeit als für L und L' gleich [1]) betrachtet werden dürfen) die Eindrückung η' mit ziemlicher Sicherheit zu bestimmen. Durch $\frac{1}{2}(\eta_1 + \eta_2) - \eta' = \eta$ findet sich die wahre Biegung, die, in die Formel (5) eingesetzt, die Ableitung des Biegungsmoduls s'_{33} gestattet. Letztere Größe ist in den Tabellen, für die beiden Stäbchenarten mit s'_g und s'_w bezeichnet, am Schluß der zweiten Zeile aufgeführt.

II. Biegungen.

G. No. 1.	$D = 1124{,}4$	$B = 5942$	$L = 44{,}2$	$P = 40$
	$\eta_1 = 162{,}6$	$\eta_2 = 161{,}6$	$\eta' = 1{,}2$	$s'_g = 47{,}8_5 . 10^{-10}$
G. No. 2.	$G = 1123{,}5$	$B = 5961$	$L = 64{,}2$	$P = 20$
	$\eta_1 = 243{,}2$	$\eta_2 = 243{,}7$	$\eta' = 0{,}8$	$s'_g = 47{,}2 . 10^{-10}$
G. No. 3.	$D = 1131{,}8$	$B = 5960$	$L = 54{,}2$	$P = 20$
	$\eta_1 = 144{,}3$	$\eta_2 = 143{,}5$	$\eta' = 0{,}8$	$s'_g = 47{,}35 . 10^{-10}$
		dasselbe mit $P = 25$.		
	$\eta = 179{,}9$	$\eta_2 = 181{,}9$	$\eta' = 1{,}0$	$s'_g = 47{,}5 . 10^{-10}$
G. No. 4.	$D = 1135{,}8$	$B = 5958$	$L = 48{,}2$	$P = 35$
	$\eta_1 = 175{,}5$	$\eta_2 = 177{,}9$	$\eta' = 1{,}1$	$s'_g = 47{,}6 . 10^{-10}$

1) Die Eindrückung variiert außer mit der Orientierung des Stäbchens auch mit der Krümmung des unterstützten Teiles der Querschnittkurve.

G. No. 5.	$D = 1132{,}9$	$B = 5971$	$L = 46{,}2$	$P = 40$
	$\eta_1 = 177{,}2$	$\eta_2 = 179{,}0$	$\eta' = 1{,}2$	$s'_g = 47{,}4.10^{-10}$
G. No. 6.	$D = 1125{,}6$	$B = 5953$	$L = 40{,}2$	$P = 40$
	$\eta_1 = 122{,}3$	$\eta_2 = 122{,}5$	$\eta' = 1{,}2$	$s'_g = 48{,}2.10^{-10}$
			Mittelwert	$s'_g = 47{,}6_1.10^{-10}$
W. No. 1.	$D = 1232{,}0$	$B = 5764$	$L = 40{,}2$	$P = 40$
	$\eta_1 = 107{,}0$	$\eta_2 = 109{,}0$	$\eta' = 2{,}2$	$s'_w = 53{,}0.10^{-10}$
W. No. 2.	$D = 1285{,}3$	$B = 5521$	$L = 38{,}2$	$P = 40$
	$\eta_1 = 84{,}1_5$	$\eta_2 = 85{,}3_5$	$\eta' = 2{,}2$	$s'_w = 52{,}8.10^{-10}$
W. No. 3.	$D = 728{,}2$	$B = 6110$	$L = 38{,}2$	$P = 40$
	$\eta_1 = 413{,}1$	$\eta_2 = 415{,}7$	$\eta' = 4{,}1$	$s'_w = 52{,}8.10^{-10}$
W. No. 4.	$D = 913{,}9$	$B = 5550$	$L = 34{,}2$	$P = 40$
	$\eta_1 = 169{,}6$	$\eta_2 = 167{,}9$	$\eta' = 4{,}1$	$s'_w = 53{,}0.10^{-10}$
W. No. 5.	$D = 740{,}5$	$B = 6120$	$L = 26{,}2$	$P = 40$
	$\eta_1 = 133{,}1$	$\eta_2 = 132{,}5$	$\eta' = 4{,}1$	$s'_w = 54{,}0.10^{-10}$
			Mittelwert	$s'_w = 53{,}2.10^{-10}$

Diese Zahlen, von denen die s'_g gemäß dem besseren Material die größere Übereinstimmung zeigen, setzen als Druckeinheit gr pro cm^2 voraus.

Herr Beckenkamp[1]) hat die reziproken Moduln $1/s'_w$ und $1/s'_g$, die als Biegungs**widerstände** bezeichnet werden können, mit dem Warburg-Kochschen Apparate gemessen und seine Resultate sind in den hier benutzten Einheiten

$$1/s'_w = 1{,}79.10^8,\ 1/s'_g = 2{,}00.10^8,$$

während meine Zahlen

$$1/s'_w = 1{,}88.10^8,\ 1/s'_g = 2{,}10.10^8$$

ergeben. Die ersteren Zahlen sind um ca. 5 % kleiner, als die von mir gefundenen. Man wird **zu klein** gefundene Biegungswiderstände im allgemeinen für wahrscheinlicher halten als **zu groß** gefundene; in der Tat wirkt die wohl wichtigste Fehlerquelle der Eindrückung an den Unterstützungsstellen in der Richtung, die Biegungswiderstände zu klein erscheinen zu lassen. Ich habe Bedenken, ob die von Herrn Beckenkamp benutzte Methode, die betreffende (bei den von ihm beobachteten kurzen Stäbchen sehr wesentliche) Fehlerquelle vollständig erledigt. Bei meinen Beobachtungen kommt sie infolge der beträchtlichen Länge der verfügbaren Präparate von vorn herein weit weniger zur Geltung.

6. Die Resultate der Drillungsbeobachtungen finden sich in der nachstehenden Tabelle in ähnlicher Anordnung, wie sie die vorige Tabelle den Ergebnissen der Biegungsbeobachtungen gegeben hat.

1) J. Beckenkamp, Zeitschr. für Krist. u. Min. **10**, 41, 1885.

Die Längen L wurden durch Messung des Abstandes der Marken bestimmt, welche die Spiegelklemmen in vier an das Stäbchen gekitteten kurzen Staniolstreifchen eingedrückt hatten. Auf die Zusammenstellung der Dimensionen folgt die Angabe der bei den Drillungen an der Millimeterskala gemachten Ablesungen σ, u. zw. wenn die Belastung erst durch die Wagschale S allein, dann bei aufgelegten n und $2n$ Gramm gemacht wurde. Diese Beobachtungen sind in der oben beschriebenen Weise wegen der Achsenreibung im belasteten Zustande korrigiert; sie enthalten aber die Wirkung dieser Reibung noch im unbelasteten Zustande. Die angegebenen Zahlen für die Drehungen an der rechten oder der linken Rolle des Apparates (rR und lR) sind somit nicht unmittelbar vergleichbar; letzteres gilt aber für die Differenzen der auf S, $S+p$, $S+2p$ bezüglichen Zahlen.

Über die Berechnung des Drillungsmoduls s_g'' wird im Anschluß an die Tabellen, gemäß dem S. 88 allgemein zur Theorie Mitgeteilten berichtet werden.

III. Drillungen.

G. No. 1.	$D = 1131{,}0$		$B = 5942$		$L = 30{,}26$
	S	$S+10$	$S+20$		
$lR.$	$\sigma = 29{,}8$	82,7	136,2		
$rR.$	30,0	82,6	135,8	$\sigma_{20} = 106{,}3$	$s_g'' = 120{,}7 . 10^{-10}$
G. No. 2.	$D = 1123{,}9$		$B = 5965$		$L = 45{,}17$
	S	$S+10$	$S+20$		
$rR.$	$\sigma = 47{,}5$	129,6	212	$\sigma_{20} = 164{,}4$	$s_g'' = 123{,}3 . 10^{-10}$
G. No. 3.	$D = 1132{,}7$		$B = 5964$		$L = 38{,}15$
	S	$S+10$	$S+20$		
$rR.$	$\sigma = 36{,}9$	103,0	170,5		
$lR.$	38,8	106,3	173,7	$\sigma_{20} = 134{,}4$	
$lR.$	$\sigma = 37{,}6$	104,6	172,5		
$rR.$	37,5	104,3	171,5	$\sigma_{20} = 134{,}5$	$s_g'' = 122{,}1 . 10^{-10}$
G. No. 4.	$D = 1147{,}0$		$B = 5958$		$L = 36{,}83$
	S	$S+15$	$S+30$		
$rR.$	$\sigma = 33{,}9$	129,9	226,0	$\sigma_{30} = 192{,}2$	$s_g'' = 124{,}4 . 10^{-19}$
G. No. 5.	$D = 1133{,}1$		$B = 5974$		$L = 32{,}71$
	S	$S+15$	$S+30$		
$lR.$	$\sigma = 31{,}2$	117,4	204,7		
$rR.$	32,9	119,7	206,6	$\sigma_{30} = 173{,}8$	$s_g'' = 123{,}0 . 10^{-10}$.

7. Aus den Zahlen von Tabelle II ergibt sich

$$s_w' = s_{11} = 53{,}2 . 10^{-10},$$
$$s_g' = \tfrac{1}{4}(2(s_{11} + s_{23}) + s_{44}) = 47{,}6 . 10^{-10},$$

und wenn man in der Formel (9) zunächst das der Einheit jedenfalls ziemlich nahe Verhältnis s'_{44}/s'_{55} mit Eins vertauscht, liefert III in erster Annäherung als Mittelwert

$$s''_g = s'_{44} = \tfrac{2}{3}(s_{11} - s_{12} + s_{44}) \sim 122{,}3 . 10^{-10}.$$

Hieraus folgen die vorläufigen Werte

$$s_{23} \sim -15{,}2 . 10^{-10}, \quad s_{44} \sim 114{,}8 . 10^{-10},$$

und hiermit nach (11)

$$s'_{55} \sim 129{,}7 . 10^{-10}, \quad s'_{45} \sim 10{,}6 . 10^{-10}.$$

In der Formel (9) ist nach S. 4 bei dem zweiten Glied in der Klammer des Nenners $\frac{1}{2}(s'^2_{45}/(s'_{44}s'_{55}))^2$ neben Eins vernachlässigt. Dieser Quotient ist hier gleich $\frac{1}{2}(1{,}12/149)^2$, somit also völlig unmerklich, die in (9) enthaltene Annäherung also weitaus genügend. Rechnet man nun die Drillungsbeobachtungen mit den obigen vorläufigen Werten von s'_{44} und s'_{55}, d. h. $\sqrt{s'_{44}/s'_{55}} = 0{,}971$ im Nenner von (9) aus, so kommt man zu einem mittleren Werte

$$s'_{44} = 122{,}7 . 10^{-10},$$

woraus dann

$$s_{44} = 115{,}2 . 10^{-10}, \quad -s_{23} = 15{,}6 . 10^{-10}$$

folgt. Für s'_{55} ergibt sich $130{,}1 . 10^{-10}$ und für $\sqrt{s'_{44}/s'_{55}}$ sehr nahe wiederum 0,971. Mit dem letzteren Wert sind die in der obigen Tabelle III aufgeführten Zahlen für s''_g berechnet. Die zuletzt angegebenen Werte für s_{44} und s_{23} sind demgemäß als die definitiven zu betrachten.

Es mag bemerkt werden, daß die Drillung eines der (nach Obigem nicht sehr hochwertigen) Stäbchen W das Ergebnis lieferte:

W. No. 4. $D = 914$ $B = 5552$ $L = 26{,}82$

		S	$S+10$	$S+20$		
rR.	$\sigma =$	48,1	135,4	224,5		
lR.		49,8	137,6	226,4	$\sigma_{20} = 176{,}5$	$s''_w = 113{,}0 . 10^{-10}$.

Da nun $s''_w \equiv s_{44}$ ist, so gibt diese Beobachtung dafür einen Wert, der mit dem aus allen G-Beobachtungen folgenden $115{,}2 . 10^{-10}$ befriedigend übereinstimmt.

Die gewonnenen Zahlen

13) $s_{11} = 53{,}2 . 10^{-10}, \quad s_{23} = -15{,}6 . 10^{-10}, \quad s_{44} = 115{,}2 . 10^{-10}$

lassen direkt erkennen, daß die elastische Anisotropie von Alaun relativ gering ist. Für einen isotropen Körper gilt nämlich die Beziehung $2(s_{11} - s_{23}) = s_{44}$, und man sieht, daß die Abweichung von derselben für Alaun ziemlich klein ist. Nebenbei sind die

Absolutwerte der Moduln für diesen Kristall auffallend groß; für Steinsalz gilt z. B.

$$s_{11} = 23{,}8 \,.\, 10^{-10}, \; s_{23} = -5{,}17 \,.\, 10^{-10}, \; s_{44} = 77{,}3 \,.\, 10^{-10}.$$

Die elastischen Widerstände von Alaun sind demgemäß klein, wie das die Zahlen für die Hauptelastizitätskonstanten nach andrer Richtung bestätigen. Die Rechnung nach den Formeln

$$14) \quad c_{11} = \frac{s_{11} + s_{23}}{(s_{11} - s_{23})(s_{11} + 2\,s_{23})}, \; c_{23} = \frac{-s_{23}}{(s_{11} - s_{23})(s_{11} + 2\,s_{23})}, \; c_{44} = \frac{1}{s_{44}}$$

liefert die Werte

$$15) \quad c_{11} = 2{,}48 \,.\, 10^{8}, \; c_{23} = 1{,}03 \,.\, 10^{8}, \; c_{44} = 0{,}86 \,.\, 10^{8}.$$

Die relativ geringe Verschiedenheit von c_{23} und c_{44} deutet auf das Vorhandensein geringer Polarität bei den Molekularkräften; bei deren völligem Fehlen gilt ja die Cauchysche Relation $c_{23} = c_{44}$. Zur Vergleichung mit den obigen Zahlen seien wieder die entsprechenden für Steinsalz hierhergesetzt. Dort gilt

$$c_{11} = 4{,}77 \,.\, 10^{8}, \; c_{23} = 1{,}32 \,.\, 10^{8}, \; c_{44} = 1{,}29 \,.\, 10^{8},$$

also Werte, die den obigen nahe liegen, übrigens der Cauchyschen Relation fast genau entsprechen.

Anhang. Die Berechnung der Elastizitätsmoduln s'_{hk} für ein beliebiges Koordinatensystem $X'\,Y'\,Z'$ aus den auf das Hauptachsensystem XYZ des (zunächst beliebigen) Kristalles bezüglichen Hauptmoduln s_{hk} ist im allgemeinen recht umständlich. Ich habe an einer anderen Stelle[1]) darauf aufmerksam gemacht, daß man die Rechnung vereinfachen kann durch Berücksichtigung der Tatsache, daß sich gewisse Aggregate von Moduln wie gewisse Produkte von Koordinaten transformieren. Es gibt darüber die folgende kleine Tabelle Auskunft:

$$16) \quad \begin{array}{cccccc} s_{11}, \ldots & \frac{1}{3}(s_{23} + \frac{1}{2}s_{44}), \ldots & \frac{1}{6}(s_{14} + s_{56}), \ldots & \frac{1}{2}s_{15}, \ldots & \frac{1}{2}s_{16}, \ldots \\ x^4, \ldots & y^2 z^2, \quad \ldots & x^2 yz, \quad \ldots & x^3 z, \ldots & x^3 y, \ldots \end{array}$$

$$\begin{array}{ccc} 4\,s_{23} - s_{44}, \ldots & s_{56} - 2\,s_{14}, \ldots \\ x^2, \quad \ldots & yz, \quad \ldots \end{array}$$

Dabei sind unter die Aggregate von Moduln die Produkte von Koordinaten gestellt, die sich jenen gleich transformieren, und die Punkte bezeichnen Ausdrücke, die aus den hingeschriebenen durch die zyklischen Vertauschungen (1, 2, 3), (4, 5, 6) und (x, y, z) entstehen.

---

1) W. Voigt, Kristallphysik p. 579 u. f.

Die Aggregate der zweiten Reihe sind hiernach direkt Tensor-, die der ersten Bitensorkomponenten.

Für das reguläre System kann man aber die Berechnung der Moduln s'_{hk} durch gewisse Überlegungen noch weiter vereinfachen. Für dieses System muß ein jeder Material-Tensor Kugelcharakter haben, d. h., seine Komponenten erster Art müssen für alle Koordinatensysteme einander gleich sein, die zweiter Art müssen verschwinden. Dies gibt zunächst die ganz allgemeinen Beziehungen

$$17)\qquad \begin{gathered} 4s'_{23}-s'_{44} = 4s'_{31}-s'_{55} = 4s'_{12}-s'_{66} = 4s_{12}-s_{44}, \\ s_{56}-2s'_{14} = s'_{64}-2s'_{25} = s'_{45}-2s'_{36} = 0. \end{gathered}$$

Es kann demnach nun auch in der ersten Zeile der obigen Tabelle $\frac{1}{6}(s_{14}+s_{56}), ..$ durch $\frac{1}{2}s_{14}, ..$ ersetzt werden.

Aus der ersten Zeile der Tabelle (16) ergibt sich hiernach, daß allgemein gilt:

$$18)\qquad \begin{gathered} s_{11}+\tfrac{1}{3}(s_{12}+\tfrac{1}{2}s_{66})+\tfrac{1}{3}(s_{31}+\tfrac{1}{2}s_{55}), .. \\ (x^2+y^2+z^2)x^2, .. \\ \tfrac{1}{2}s_{14}+\tfrac{1}{2}s_{24}+\tfrac{1}{2}s_{34}, .. \\ (x^2+y^2+z^2)yz, .. \end{gathered}$$

Diese Aggregate sind also gleichfalls Tensorkomponenten und müssen für das reguläre System und jedes Koordinatenkreuz X' Y' Z' den Bedingungen genügen

$$19)\qquad \begin{gathered} s'_{11}+\tfrac{1}{3}(s'_{12}+s'_{33})+\tfrac{1}{6}(s'_{11}+s'_{55}) = .. = s'_{11}+\tfrac{2}{3}s'_{23}+\tfrac{1}{3}s'_{44}; \\ s'_{14}+s'_{24}+s'_{34} = .. = 0. \end{gathered}$$

Weiter kommt in Betracht, daß für das reguläre System alle die Hauptmoduln $s_{14}, .. s_{15}, .. s_{16}, ..$ verschwinden, so daß von der ersten Reihe in (16) nur die sechs Aggregate $s_{11}, .., \frac{1}{3}(s_{23}+\frac{1}{2}s_{44}), ..$ übrig bleiben, die sich wie $x^4, ...$ und $y^2z^2, ..$ verhalten. Will man also irgend einen der Moduln s'_{hk} berechnen, der sich wie $x^\alpha . y^\beta . z^\gamma$ bei $\alpha+\beta+\gamma = 4$ transformiert, so braucht man in dem Produkt $x'^\alpha . y'^\beta . z'^\gamma$ nur die mit $x^4, ..$ und $y^2z^2, ..$ proportionalen Glieder zu bilden. Das geht am leichtesten, wenn man $x'^\alpha . y'^\beta . z'^\gamma$ in zwei Faktoren zweiten Grades zerlegt, z. B. $x'y'z'^2$ in $(x'.y').z'^2$ und diese beiden Faktoren in x, y, z ausdrückt.

Benutzt man das System der Transformationskoeffizienten

20)

	x'	y'	z'
x	α_1	β_1	γ_1
y	α_2	β_2	γ_2
z	α_3	β_3	γ_3

sowie, daß hiernach

$$21)\quad \begin{aligned} x'y' &= \alpha_1\beta_1 x^2 + \cdot\cdot + (\alpha_2\beta_3 + \beta_2\alpha_3) yz + \cdot\cdot, \\ z'^2 &= \gamma_1^2 x^2 + \cdot\cdot + 2\gamma_2\gamma_3 yz + \cdot\cdot, \end{aligned}$$

so erhält man für $\frac{1}{2} s_{36} = \frac{1}{4} s_{45}$, was sich transformiert wie xyz^2, direkt den Ausdruck

$$22)\quad \begin{aligned} \tfrac{1}{2} s'_{36} &= s_{11}(\alpha_1\beta_1\gamma_1^2 + \cdot\cdot) \\ &+ \tfrac{1}{3}(s_{12} + \tfrac{1}{2} s_{44})[\alpha_1\beta_1(\gamma_2^2 + \gamma_3^2) + \cdot\cdot + 2\gamma_2\gamma_3(\alpha_2\beta_3 + \beta_2\alpha_3) + \cdot\cdot]. \end{aligned}$$

Nun folgt aber aus

$$(\alpha_1\gamma_1 + \cdot\cdot)(\beta_1\gamma_1 + \cdot\cdot) = 0 \text{ die Beziehung}$$

$$23)\quad \alpha_1\beta_1\gamma_1^2 + \cdot\cdot + \gamma_2\gamma_3(\alpha_2\beta_3 + \alpha_3\beta_2) + \cdot\cdot = 0,$$

und deren Berücksichtigung liefert aus Obigem den Ausdruck:

$$24)\quad \tfrac{1}{2} s'_{36} = \tfrac{1}{4} s'_{45} = (s_{11} - s_{23} - \tfrac{1}{2} s_{44})(\alpha_1\beta_1\gamma_1^2 + \alpha_2\beta_2\gamma_2^2 + \alpha_3\beta_3\gamma_3^2).$$

Ähnlich gewinnt man aus

$$25)\quad \begin{aligned} y'^2 &= \beta_1^2 x^2 + \cdot\cdot + 2\beta_2\beta_3 yz + \cdot\cdot \\ z'^2 &= \gamma^1 x^2 + \cdot\cdot + 2\gamma_2\gamma_3 yz + \cdot\cdot \end{aligned}$$

für das mit $y^2 z^2$ parallelgehende Aggregat $\frac{1}{3}(s_{23} + \frac{1}{2} s_{44})$ unmittelbar

$$26)\quad \begin{aligned} \tfrac{1}{3}(s'_{23} + \tfrac{1}{2} s'_{44}) &= s_{11}(\beta_1^2\gamma_1^2 + \cdot\cdot) \\ &+ \tfrac{1}{3}(s_{23} + \tfrac{1}{2} s_{44})[\beta_1^2(\gamma_2^2 + \gamma_3^2) + \cdot\cdot + 4\beta_2\beta_3\gamma_2\gamma_3 + \cdot\cdot]. \end{aligned}$$

Nun folgt aus

$$(\beta_1\gamma_1 + \cdot\cdot)^2 = 0 \text{ die Beziehung}$$

$$27)\quad \beta_1^2\gamma_1^2 + \cdot\cdot + 2(\beta_2\beta_3\gamma_2\gamma_3 + \cdot\cdot) = 0,$$

und mit ihrer Hilfe liefert die vorige Formel

$$28)\quad \tfrac{1}{3}(s'_{23} + \tfrac{1}{2} s'_{44}) = \tfrac{1}{3}(s_{23} + \tfrac{1}{2} s_{44}) + (s_{11} - s_{23} - \tfrac{1}{2} s_{44})(\beta_1^2\gamma_1^2 + \cdot\cdot).$$

Hinzu kommt nach (17)

$$\tfrac{1}{3}(s'_{23} - \tfrac{1}{4} s'_{44}) = \tfrac{1}{3}(s_{23} - \tfrac{1}{4} s_{44})$$

was zusammen mit dem eben Erhaltenen liefert

$$29)\quad \begin{aligned} s'_{23} &= s_{23} + (s_{11} - s_{23} - \tfrac{1}{2} s_{44})(\beta_1^2\gamma_1^2 + \cdot\cdot), \\ s'_{44} &= s_{44} + 4(s_{11} - s_{23} - \tfrac{1}{2} s_{44})(\beta_1^2\gamma_1^2 + \cdot\cdot). \end{aligned}$$

Analog gilt

$$30)\quad \begin{aligned} s'_{31} &= s_{31} + (s_{11} - s_{23} - \tfrac{1}{2} s_{44})(\alpha_1^2\gamma_1^2 + \cdot\cdot) \\ s'_{55} &= s_{55} + 4(s_{11} - s_{23} - \tfrac{1}{2} s_{44})(\alpha_1^2\gamma_1^2 + \cdot\cdot). \end{aligned}$$

Für die speziell S. 5 vorausgesetzte Orientierung des $X' Y' Z'$-Achsenkreuzes erhält man somit

$$
\begin{aligned}
s'_{45} &= (2(s_{11} - s_{23}) - s_{44}) \tfrac{1}{3} \sqrt{2}, \\
s'_{44} &= s_{44} + \tfrac{2}{3}(s_{11} - s_{23} - \tfrac{1}{2} s_{44}) = \tfrac{2}{3}(s_{11} - s_{12} + s_{44}), \\
s'_{55} &= s_{44} + \tfrac{4}{3}(s_{11} - s_{23} - \tfrac{1}{2} s_{44}) = \tfrac{1}{3}(4(s_{11} - s_{12}) + s_{44}).
\end{aligned}
\tag{31}
$$

Dies sind aber die Ausdrücke, die oben angewendet worden sind.

Göttingen, Februar 1919.

Über Transformations- und Klassenpolygone.

(Vorläufige Mitteilung aus einer Herrn F. Klein zu seinem siebzigsten Geburtstage gewidmeten Arbeit.)

Von

Robert Fricke in Braunschweig.

Vorgelegt in der Sitzung vom 30. Mai 1919.

Eines der wesentlichsten Hilfsmittel, deren sich Klein bei seinen ersten Untersuchungen über die Modulargleichungen erster Stufe der elliptischen Funktionen[1]) bedient, ist der von ihm selbst als „Fundamentalpolygon" bezeichnete Bereich. Es handelt sich um den Diskontinuitätsbereich derjenigen in der Modulgruppe, d. i. in der Gruppe aller ganzzahligen Substitutionen $\omega' = \frac{\alpha\omega + \beta}{\gamma\omega + \delta}$ der Determinante 1, enthaltenen Kongruenzuntergruppe n ter Stufe, die durch $\beta \equiv 0 \pmod{n}$ erklärt ist. Diese Untergruppe hat in der Gesamtgruppe bekanntlich den Index:

$$1) \qquad \psi(n) = n\,\Pi\left(1 + \frac{1}{p}\right),$$

wo sich das Produkt auf die verschiedenen Primfaktoren des Transformationsgrades n bezieht. Übrigens benutze ich lieber die mit der genannten Untergruppe gleichberechtigte Untergruppe, die durch $\gamma \equiv 0 \pmod{n}$ erklärt ist, und nenne diese $\Gamma_{\psi(n)}$, die erste dagegen $\Gamma'_{\psi(n)}$. Die $\Gamma_{\psi(n)}$ gehört nämlich zur ersten Haupttransformation $n$ten Grades, die $\Gamma'_{\psi(n)}$ aber zur zweiten. Der Diskontinuitätsbereich der $\Gamma_{\psi(n)}$ werde durch $P_{t(n)}$ bezeichnet und das „Transformationspolygon $n$ten Grades" genannt.

1) S. den Abschn. II der Abhandlung Kleins „Über die Transformation der elliptischen Funktionen usw.", Math. Ann. Bd. 14, S. 127 ff.

Der Übergang zum Diskontinuitätsbereiche $P'_{t(n)}$ der $\Gamma'_{\psi(n)}$ ist dann sehr einfach: Man gewinnt $P'_{t(n)}$ unmittelbar aus $P_{t(n)}$, wenn man in diesen Bereich das Dreiecksnetz der Gruppe Γ nach einem auf $\frac{1}{n}$ reduzierten Maßstabe einzeichnet. Es hängt dies zusammen mit der Tatsache, daß das Polygon $P_{t(n)}$ durch die symbolisch mit W_n zu bezeichnende Substitution:

2) $$\omega' = \frac{-1}{n\omega}$$

in sich transformiert wird, wobei dann eben das Dreiecksnetz von $P_{t(n)}$ in das von $P'_{t(n)}$ übergeht.

Die Transformation W_n spielt bereits in der oben genannten Abhandlung von Klein eine bekannte grundlegende Rolle. Es ist nun die Absicht der vorliegenden Notiz, aus der Einführung der Substitution W_n zunächst noch ohne funktionentheoretische Gesichtspunkte einige neue Folgerungen zu ziehen.

Die Gruppe $\Gamma_{\psi(n)}$ wird durch Zusatz von W_n zu einer Gruppe:

3) $$\Gamma_n = \Gamma_{\psi(n)} + \Gamma_{\psi(n)} \cdot W_n$$

erweitert, in der die $\Gamma_{\psi(n)}$ eine ausgezeichnete Untergruppe des Index 2 ist. Dabei kann man die Nebengruppe $\Gamma_{\psi(n)} \cdot W_n$ erklären als bestehend aus allen Substitutionen $\omega' = \frac{\alpha\omega + \beta}{\gamma\omega + \delta}$ der Determinante n, deren ganzzahlige Koeffizienten die Bedingungen:

4) $$\alpha \equiv 0, \quad \gamma \equiv 0, \quad \delta \equiv 0 \pmod{n}$$

erfüllen[1]).

Der durch $P_{k(n)}$ zu bezeichnende Diskontinuitätsbereich der Γ_n ist nun ein Gebilde von besonderem Interesse: Er steht in unmittelbarer Beziehung zu den Klassen ganzzahliger binärer quadratischer Formen sowohl der negativen Diskriminante $D = -4n$ wie auch derjenigen der po-

---

1) Führt man $\omega\sqrt{n} = \zeta$ als Variable ein, so wird Γ_n die Gruppe aller „unimodularen" Substitutionen:

$$\zeta' = \frac{\alpha\zeta + \beta\sqrt{n}}{\gamma\sqrt{n}\,\zeta + \delta}, \qquad \zeta' = \frac{\alpha\sqrt{n}\,\zeta + \beta}{\gamma\zeta + \delta\sqrt{n}}$$

mit ganzzahligen α, β, γ, δ. Die Gruppeneigenschaft ist hier besonders einleuchtend. Die Substitution $\zeta = \omega\sqrt{n}$ führt zur $\Gamma_{\psi(n)}$ zurück, die Substitution $\zeta = \frac{\omega}{\sqrt{n}}$ zur $\Gamma'_{\psi(n)}$.

sitiven Diskriminante $D = +4n$. Der fragliche Bereich möge dieserhalb als „Klassenpolygon $n$ten Grades“ bezeichnet werden. In dieser Beziehung darf man wohl den klarsten Ausdruck für den Zusammenhang erkennen, der zwischen der Theorie der ganzzahligen binären quadratischen Formen und den elliptischen Funktionen besteht.

Zur näheren Darlegung werde zunächst gefragt, welche elliptischen und parabolischen Substitutionen in der Nebengruppe $\Gamma_{\psi(n)} \cdot W_n$ enthalten sind. Bei den elliptischen Substitutionen können nur solche der Perioden 6,4 und 2 auftreten. Eine einfache Rechnung zeigt, daß Substitutionen der Periode 6 nur bei $n = 3$, solche der Periode 4 nur bei $n = 2$ auftreten. Ebenso leicht zeigt man, daß parabolische Substitutionen nur bei $n = 4$ vorkommen. Diese drei niedersten Fälle sind aber noch so einfach, daß sie ausgeschlossen werden mögen und also $n > 4$ vorausgesetzt wird.

Eine Substitution der Periode 2 der Nebengruppe hat nun die Gestalt:

$$5)\qquad \omega' = \frac{n\alpha_0\omega + \beta}{n\gamma_0\omega - n\alpha_0}, \quad n\alpha_0^2 + \beta\gamma_0 = -1,$$

wo α_0, β, γ_0 ganze Zahlen sind. Der Fixpunkt der Substitution ist der Nullpunkt der symbolisch durch $(n\gamma_0, -2n\alpha_0, -\beta)$ zu bezeichnenden ganzzahligen binären quadratischen Form:

$$6)\qquad (n\gamma_0, -2n\alpha_0, -\beta) = n\gamma_0 x^2 - 2n\alpha_0 xy - \beta y^2$$

der Diskriminante $D = -4n$. Da β gegen $n\alpha_0$ teilerfremd ist, so ist diese Form entweder ursprünglich, oder sie hat den Teiler 2. Die letztere Möglichkeit erfordert $n \equiv 3 \pmod 4$, wo man an Stelle der Form (6) besser mit der „ursprünglichen“ Form:

$$7)\qquad \left(n\frac{\gamma_0}{2}, -n\alpha_0, -\frac{\beta}{2}\right) = n\frac{\gamma_0}{2}x^2 - n\alpha_0 xy - \frac{\beta}{2}y^2$$

der Diskriminante $D = -n$ arbeitet. Übrigens darf man sich auf „positive“ Formen beschränken, da ein gleichzeitiger Zeichenwechsel der vier Substitutionskoeffizienten ohne Folge ist.

Ist andrerseits eine beliebige Klasse ursprünglicher positiver Formen der Diskriminante $D = -4n$ oder (im besonderen Falle $n \equiv 3 \pmod 4$) der Diskriminante $D = -n$ vorgelegt, so kann man der Klasse eine Form (a', b', c) mit einem gegen n teilerfremden c entnehmen und sodann durch eine geeignete Substitution $x' = x$, $y' = \gamma x + y$ zu einer äquivalenten Form $(na_0, 2nb_0, c)$ bezw. (na_0, nb_0, c) mit ganzen Zahlen a_0, b_0 übergeben. Sind

aber zunächst für $D = -4n$ irgend zwei der Klasse angehörende Formen der Gestalt (6) in $(na_0, 2nb_0, c)$ und $(na'_0, 2nb'_0, c')$ vorgelegt, so folgt aus:

$$na'_0 = na_0\delta^2 - 2nb_0\gamma\delta + c\gamma^2,$$
$$nb'_0 = -na_0\beta\delta + nb_0(\alpha\delta + \beta\gamma) - c\alpha\gamma,$$

daß γ^2 und $\alpha\gamma$ und also auch γ durch n teilbar sind. Transformiert man andrerseits $(na_0, 2nb_0, c)$ durch eine Substitution $\begin{pmatrix}\alpha, & \beta\\ \gamma, & \delta\end{pmatrix}$ mit $\gamma \equiv 0 \pmod{n}$, so ergibt sich auch stets wieder eine Form $(na'_0, 2nb'_0, c')$ der Gestalt (6). Die in der Klasse enthaltenen $(na_0, 2nb_0, c)$ sind also gerade diejenigen Formen, die mit einer unter ihnen bezüglich der $\Gamma_{\psi(n)}$ äquivalent sind. Zu dem gleichen Ergebnisse gelangt man im besonderen Falle $n \equiv 3 \pmod 4$ auch für die in der einzelnen ursprünglichen Klasse der Diskriminante $D = -n$ enthaltenen Formen der Gestalt (na_0, nb_0, c).

Der einzelnen Form $(na_0, 2nb_0, c)$ oder (na_0, nb_0, c) entspricht nun umgekehrt auch immer eine Substitution der Periode 2:

$$8)\qquad \omega' = \frac{-nb_0\omega - c}{na_0\omega + nb_0} \quad \text{bezw.} \quad \omega' = \frac{-nb_0\omega - 2c}{2na_0\omega + nb_0}$$

in der Nebengruppe $\Gamma_{\psi(n)} \cdot W_n$. Faßt man also die innerhalb der $\Gamma_{\psi(n)}$ gleichberechtigten Substitutionen je in eine „Klasse" zusammen, so besteht der Satz: Die Klassen elliptischer Substitutionen der Periode 2 in der Nebengruppe $\Gamma_{\psi(n)} \cdot W_n$ sind umkehrbar eindeutig den Klassen ursprünglicher positiver Formen der Diskriminante $D = -4n$ bezw. der beiden Diskriminanten $D = -n$ und $D = -4n$ zugeordnet. Wird mit $h(-D)$ die Klassenanzahl ursprünglicher positiver Formen der negativen Diskriminante D bezeichnet, so kann man den Satz auch in folgender Fassung aussprechen: Beim Fortgange vom Transformationspolygone $P_{t(n)}$ zum Klassenpolygone $P_{k(n)}$ treten $h(4n)$ bezw. $h(n) + h(4n) = 4h(n)$ neue feste Polygonecken mit Winkeln π auf; sonstige elliptische Ecken oder parabolische Spitzen von $P_{k(n)}$ kommen mit den gleichen Winkeln auch schon am Transformatiospolygone vor. Übrigens kann es bei der wirklichen Zeichnung eines einzelnen Klassenpolygons zweckmäßig sein, eine neu hinzutretende Ecke gelegentlich in einen „Eckenzyklus" aufzulösen, wo dann die Summe der Winkel des Zyklus gleich π ist.

Die Gruppen $\Gamma_{\psi(n)}$ und Γ_n sind durch die Spiegelung $\omega' = -\bar{\omega}$ an der imaginären Axe[1]) auf Gruppen $\overline{\Gamma}_{\psi(n)}$ und $\overline{\Gamma}_n$ erweiterungsfähig, in denen die ursprünglichen Gruppen ausgezeichnete Untergruppen des Index 2 sind. Sind die Bereiche $P_{t(n)}$ und $P_{k(n)}$ symmetrisch bezüglich der imaginären ω-Axe gezeichnet, so stellen die etwa zur linken Seite der genannten Axe gelegenen Hälften jener Polygone Diskontinuitätsbereiche $\overline{P}_{t(n)}$ und $\overline{P}_{k(n)}$ der erweiterten Gruppen dar.

Die in $\overline{\Gamma}_{\psi(n)}$ enthaltenen Spiegelungen haben Symmetriekreise, die auch schon im Dreiecksnetze der gesamten Modulgruppe auftreten. Wir fragen nun weiter nach den in der Nebengruppe $\overline{\Gamma}_{\psi(n)} \cdot W_n$ enthaltenen Spiegelungen, die also beim Übergange von $\overline{P}_{t(n)}$ zu $\overline{P}_{k(n)}$ neu hinzutreten. Es sind dies die Substitutionen:

9) $$\omega' = \frac{n\alpha_0\bar{\omega} - \beta}{n\gamma_0\bar{\omega} - n\alpha_0}, \qquad n\alpha_0^2 - \beta\gamma_0 = 1,$$

deren Symmetriekreise gegeben sind durch:

10) $$n\gamma_0(\xi^2 + \eta^2) - 2n\alpha_0\xi + \beta = 0.$$

Der einzelne Kreis (10) wird zur geometrischen Deutung für zwei nur im Vorzeichen verschiedene ganzzahlige binäre quadratische Formen $(\pm n\gamma_0, \mp 2n\alpha_0, \pm\beta)$ der positiven Diskriminante $D = 4n$ benutzt. Diese beiden Formen sind stets und nur dann äquivalent, wenn der Kreis (10) durch Eckpunkte des Dreiecksnetzes der ω-Halbebene hindurchläuft, die bezüglich der Gesamtgruppe mit dem Punkte $\omega = i$ äquivalent sind. Es sollen weiterhin auch dann, wenn dies nicht der Fall ist, die beiden Klassen der Formen $(\pm n\gamma_0, \mp 2n\alpha_0, \pm\beta)$ als nicht verschieden angesehen werden. Die Form $(n\gamma_0, -2n\alpha_0, \beta)$ erweist sich wieder als ursprünglich oder als eine solche vom Teiler 2, wobei die letztere Möglichkeit jedoch nur für $n \equiv 1 \pmod 4$ eintreten kann. In diesem Falle soll an die Stelle von $(n\gamma_0, -2n\alpha_0, \beta)$ die ursprüngliche Form $\left(n\frac{\gamma_0}{2}, -n\alpha_0, \frac{\beta}{2}\right)$ der Diskriminante $D = n$ treten.

Ist umgekehrt eine beliebige ursprüngliche Formklasse der Diskriminante $D = 4n$ bezw. $D = n$ vorgelegt, so können wir erstlich wieder der Klasse eine Form (a', b', c) mit einem gegen n teilerfremden c entnehmen. Ist $D = 4n$ und also b' gerade, so kann man durch eine geeignete Substitution $x' = x$, $y' = \gamma x + y$ zu einer Form der Gestalt $(na_0, 2nb_0, c)$ übergehen. Ist aber

1) Mit $\bar{\omega}$ ist der zu $\omega = \xi + i\eta$ konjugiert komplexe Wert $\bar{\omega} = \xi - i\eta$ gemeint.

$n \equiv 1 \pmod 4$ und $D = n$, so kann man entsprechend zu einer äquivalenten Form (na_0, nb_0, c) gelangen. Auch ergibt sich wieder wie oben, daß die in der einzelnen Klasse enthaltenen Formen der Gestalt $(na_0, 2nb_0, c)$ bezw. (na_0, nb_0, c) gerade genau alle diejenigen sind, die mit einer unter ihnen bezüglich der $\Gamma_{\psi(n)}$ äquivalent sind.

Endlich entspricht umgekehrt auch wieder jeder Form $(na_0, 2nb_0, c)$ und (na_0, nb_0, c) eine Spiegelung:

11) $$\omega' = \frac{-nb_0\bar{\omega} - c}{na_0\bar{\omega} + nb_0} \quad \text{bezw.} \quad \omega' = \frac{-nb_0\bar{\omega} - 2c}{2na_0\bar{\omega} + nb_0}.$$

Faßt man also die bezüglich der $\Gamma_{\psi(n)}$ äquivalenten Symmetriekreise (gleichberechtigten Spiegelungen) in eine „Klasse" zusammen und bezeichnet mit $h'(D)$ die nach der oben festgesetzten Vorschrift abgezählte Anzahl der ursprünglichen Formklassen der positiven Diskriminante D, so besteht der Satz: Die Klassen der Spiegelungen in der Nebengruppe $\bar{\Gamma}_{\psi(n)} \cdot W_n$ sind für $n \not\equiv 1 \pmod 4$ umkehrbar eindeutig den $h'(4n)$ ursprünglichen Formklassen der Diskriminante $D = 4n$ zugeordnet, für $n \equiv 1 \pmod 4$ entsprechen sie umkehrbar eindeutig den $h'(n) + h'(4n)$ ursprünglichen Formklassen der beiden Diskriminanten $D = n$ und $D = 4n$. Man kann dem Satze auch wieder eine geometrische Fassung verleihen: Beim Fortgange vom Transformationspolygone zum Klassenpolygone treten $h'(4n)$ bezw. $h'(n) + h'(4n)$ neue Symmetriekreise als Randkurven von $P_{k(n)}$ auf, deren zugehörige Formen als Repräsentanten der $h'(4n)$ bezw. $h'(n) + h'(4n)$ fraglichen Klassen gelten können[1]. Wie übrigens oben ein einzelner elliptischer Eckpunkt von $P_{k(n)}$, um diesem Polygone eine zweckmäßige Gestalt zu verleihen, in einen Eckenzyklus aufgelöst werden konnte, so kann auch am Rande von $P_{k(n)}$ der einzelne Symmetriekreis in einen „Seitenzyklus" des Klassenpolygons aufgelöst erscheinen.

Für die Durchführung der Transformation der elliptischen Funktionen (erster Stufe) ist die Gewinnung des Klassenpolygons

1) Wählt man $P_{k(n)}$ als „Normalpolygon" mit dem „Zentrum" in der nach der Stelle $\omega = i\infty$ ziehenden Polygonspitze, so ist die Gestalt des Klassenpolygons in jedem Falle n eindeutig bestimmt. An diese Gestalten der $P_{k(n)}$ könnte man Reduktionsbedingungen für die Formen anknüpfen. Bemerkenswert ist dabei, daß die Reduktionstheorien in den beiden Fällen $D < 0$ und $D > 0$ übereinstimmende Gestalten gewinnen; bei $D > 0$ fällt die umständliche Gaußsche Theorie der vielgliedrigen Ketten reduzierter Formen fort.

von wesentlicher Bedeutung. Eine allgemeine Formel für das Geschlecht p_t des Transformationspolygons ist seit lange bekannt[1]. Für das Geschlecht des Klassenpolygons, das mit p_k bezeichnet werden mag, gilt die Regel:

12) $$p_k = \tfrac{1}{2} p_t + \tfrac{1}{2} - \tfrac{1}{4} \varepsilon_n h(4n),$$

won ε_n gleich $\frac{4}{3}$ für $n \equiv 3 \pmod 8$, gleich 2 für $n \equiv 7 \pmod 8$ und sonst gleich 1 ist. Es zeigt sich, daß über die bekannten Fälle mit $p_t = 0$ hinaus noch für die folgenden 22 Transformationsgrade:

$$n = 11,\ 14,\ 15,\ 17,\ 19,\ 20,\ 21,\ 23,\ 24,\ 26,\ 27,\ 29,\ 31,\ 32,\ 35,\ 36,\ 39,\ 41,\ 47,\ 49,\ 50,\ 71$$

das Geschlecht p_k des Klassenpolygons gleich Null ist. Für die Mehrzahl dieser Fälle habe ich in dem im Ms. nahezu beendeten zweiten Bande meines Buches über die elliptischen Funktionen die Klassenpolygone gezeichnet und die Transformationen mit Benutzung formentheoretischer Methoden durchgeführt.

Braunschweig, den 21. Mai 1919.

---

1) Man findet die Berechnung von p_t bei Gierster in den Mathem. Annalen, Bd. 14, S. 538.

Tragflügeltheorie.

II. Mitteilung.

Von

L. Prandtl.

Vorgelegt in der Sitzung vom 21. Februar 1919.

A. Mehrdeckertheorie.

1. Als Vorarbeit für die Durchbildung der Theorie des Mehrdeckers auf der Grundlage der in der ersten Mitteilung gegebenen Ansätze für die Flügeltheorie der ersten Ordnung sei zunächst die Aufgabe behandelt, von der Störungsgeschwindigkeit, die von einer „tragenden Linie" mit gegebener Auftriebsverteilung herrührt, an einem Aufpunkt A die Komponente nach einer gegebenen Richtung zu ermitteln (in den Anwendungen wird es sich hauptsächlich um die Richtung senkrecht zu einer anderen den Aufpunkt enthaltenden tragenden Linie und senkrecht zur Flugrichtung handeln).

Die Aufgabe werde zunächst auf Grund der in Nr. 12 der I. Mitteilung angegebenen Beziehungen für den Fall gelöst, daß sowohl die tragende Linie, wie auch der Aufpunkt und die vorgegebene Richtung in ein und derselben Ebene senkrecht zur Flugrichtung (einer „Querebene") enthalten seien. Dann ist die Wirkung der Querwirbel gleich Null, und die Längswirbel ergeben (vgl. Abb. 1 auf folgender Seite)

$$w = -\frac{1}{4\pi}\int_0^l \frac{d\Gamma}{ds}\frac{\sin\beta}{a}\,ds. \tag{28}$$

Da an den Flügelenden $\Gamma = 0$ ist, läßt sich durch partielle Integration hierfür schreiben:

$$w = \frac{1}{4\pi}\int_0^l \Gamma \frac{d}{ds}\left(\frac{\sin\beta}{a}\right) ds.$$

Wählt man den Aufpunkt zum Koordinatenursprung und die vorgegebene Richtung zur Z-Achse, so wird $\sin\beta/a = x/a^2$, ferner

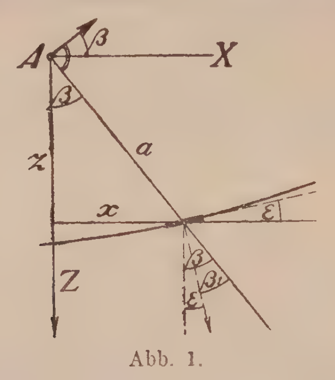


Abb. 1.

$\frac{d}{ds} = \cos\varepsilon \frac{\partial}{\partial x} - \sin\varepsilon \frac{\partial}{\partial z}$; hiermit ergibt sich nach kurzer Rechnung $\frac{d}{ds}\left(\frac{\sin\beta}{a}\right) = \frac{1}{a^2}\cos(2\beta - \varepsilon)$ und also

$$w = \frac{1}{4\pi}\int_0^l \frac{\Gamma\cos(2\beta - \varepsilon)}{a^2} ds. \tag{29}$$

Mit Hilfe dieser Gleichung läßt sich sofort der Widerstand anschreiben, der dadurch entsteht, daß ein Flügel 1 unter dem Einfluß der Störung steht, die von einem in derselben „Querebene" befindlichen Flügel 2 ausgeht. Die Störungsgeschwindigkeit für einen auf dem Flügel 1 befindlichen Aufpunkt heiße w_{12}; der Widerstand wird dann gemäß den Darlegungen in Nr. 12 der I. Mitteilung

$$W_{12} = \varrho \int \Gamma_1 ds_1 . w_{12}$$

oder wenn w_{12} nach Gl. (29) eingesetzt wird, wobei gemäß Abb. 6 β_1 anstelle von $\beta - \varepsilon$, ferner β_2 anstelle von β geschrieben wird:

$$W_{12} = \frac{\varrho}{4\pi}\int_0^{l_1}\int_0^{l_2} \frac{\Gamma_1 \Gamma_2 \cos(\beta_1 + \beta_2) ds_1 ds_2}{a^2}. \tag{30}$$

Das Doppelintegral ist in den Größen von Flügel 1 und 2 vollständig symmetrisch gebaut; wir schließen daraus, daß der Widerstand, der am Flügel 2 durch die Gegenwart des Flügels 1 entsteht, von derselben Größe sein muß, wie der hier berechnete, daß also

$$W_{12} = W_{21}$$

ist. Diese Gegenseitigkeitsbeziehung ist für die Folge von großer Wichtigkeit. Sie ist von meinem Mitarbeiter Dr. M. Munk auf einem andern Wege gefunden worden. Da es offenbar nicht darauf ankommt, daß die zusammengefaßten tragenden Elemente jeweils zu einem einzigen Flügel gehören, läßt sich der Satz auch so aussprechen:

„Greift man aus einem tragenden System, dessen Elemente sämtlich in einer „Querebene" enthalten sind, zwei Gruppen beliebig heraus, so ist derjenige Widerstandsanteil, den die Gruppe 1 durch das Geschwindigkeitsfeld der Gruppe 2 erfährt, ebenso groß, wie derjenige von Gruppe 2 im Feld von Gruppe 1".

Anmerkungen: 1) Der gegenseitige Widerstand zweier untereinander befindlichen Flügel ist positiv, der von zwei nebeneinander befindlichen Flügeln dagegen negativ; durch erstere Anordnung wird also der Gesamtwiderstand gegenüber dem Zustand weit voneinander entfernter Flügel vermehrt, durch letztere vermindert.

2) Die bei der Ableitung von Gl. (29) benutzte partielle Integration läßt eine anschauliche Deutung zu: die Geschwindigkeit w erscheint hierin aufgebaut aus den Beiträgen von lauter infinitesimalen Flügeln von der Länge ds und der Zirkulation Γ. Der Integrand von Gl. (29) stimmt in der Tat, wie man schon aus der Ableitung entnehmen kann, überein mit der Geschwindigkeit, die von zwei im Abstande ds befindlichen Wirbellinien von entgegengesetzt gleicher Wirbelstärke $= \Gamma$ hervorgebracht wird. Das Doppelintegral in Gl. (30) darf in diesem Sinn aufgefaßt werden als die Summe der Wirkungen der Wirbelschweife aller tragenden Elemente ds_2 auf alle tragenden Elemente ds_1.

3) Die Gleichungen (29) und (30) verlieren ihre Anwendbarkeit, wenn der Wert $a = 0$ vorkommt, da sie in diesem Fall die Form $\infty - \infty$ annehmen. Sie sind deshalb auch ungeeignet zur Berechnung der Geschwindigkeit w_n auf dem Flügel selbst und zu der des Eigenwiderstandes eines Flügels. Für diese Größen steht jedoch die auch sonst verwendbare Gl. (28) zur Verfügung, wobei hier von dem Integral der Hauptwert zu nehmen ist. Grundsätzlich ist allerdings auch eine Ermittlung mit Gl. (29) und (30) auf dem Wege denkbar, daß die gegenseitige Einwirkung von zwei sehr benachbarten Flügeln gleicher Art berechnet und im Endresultat zum Grenzfall des Zusammenfallens der beiden Flügel übergegangen wird. Aus einer entsprechenden Überlegung folgt u. a., daß

die Beziehung $W_{12} = W_{21}$ auch für den Grenzfall zweier räumlich zusammenfallender, im übrigen aber beliebiger tragender Systeme gilt. Hiervon wird in Nr. 5 Gebrauch gemacht werden.

2. Zur Untersuchung des allgemeinen Falls werde zunächst das Feld berechnet, das ein tragendes Element von der Länge ds mitsamt dem abgehenden Wirbelpaar in irgendeinem Raumpunkt hervorbringt. Der Koordinatenursprung werde in das tragende Element gelegt, die Richtung des Elements werde zur X-Achse gemacht; dann ist zunächst die von einem einzigen abgehenden Wirbel erzeugte Geschwindigkeit, mit $a^2 = x^2 + z^2$, $r^2 = a^2 + y^2$ und $r'^2 = a^2 + (y - y')^2$

$$\frac{\Gamma}{4\pi}\int_0^\infty \frac{dy' \sin\varphi}{r'^2} = \frac{\Gamma}{4\pi}\int_0^\infty \frac{dy' \,.\, a}{r'^3} = \frac{\Gamma}{4\pi a}\left(1 + \frac{y}{r}\right),$$

also ihre Komponenten:

$$u = \frac{\Gamma x}{4\pi a^2}\left(1 + \frac{y}{r}\right); \quad v = 0; \quad w = -\frac{\Gamma z}{4\pi a^2}\left(1 + \frac{y}{r}\right).$$

Das abgehende Wirbelpaar gibt

$$du_1 = -\frac{\partial u}{\partial x}ds; \quad dv_1 = 0; \quad dw_1 = -\frac{\partial w}{\partial x}ds.$$

Hierzu kommt noch die Wirkung des Querwirbels:

$$du_2 = 0; \quad dv_2 = -\frac{\Gamma ds}{4\pi}\frac{z}{r^3}; \quad dw_2 = \frac{\Gamma ds}{4\pi}\frac{y}{r^3}.$$

Nach einer kurzen Zwischenrechnung ergibt sich wenn die aus Abb. 2 ersichtlichen Winkel eingeführt werden:

$$\left.\begin{aligned} du &= \frac{\Gamma ds}{4\pi}\left\{(1 + \sin\alpha)\frac{\sin 2\beta}{a^2} + \frac{\sin\alpha \sin 2\beta}{2r^2}\right\} \\ dv &= -\frac{\Gamma ds}{4\pi}\cdot\frac{\cos\alpha\cos\beta}{r^2} \\ dw &= \frac{\Gamma ds}{4\pi}\left\{(1 + \sin\alpha)\frac{\cos 2\beta}{a^2} + \frac{\sin\alpha\cos^2\beta}{r^2}\right\} \end{aligned}\right\} \quad (31)$$

Zur Berechnung des Anstellwinkels und des Widerstandes braucht man die Geschwindigkeitskomponente w_n senkrecht zur Flugrichtung und senkrecht zur Richtung des beeinflußten tragenden Elementes. Ist dieses (1 in Abb. 3) gegen die X-Achse unter dem Winkel ε geneigt, so ist

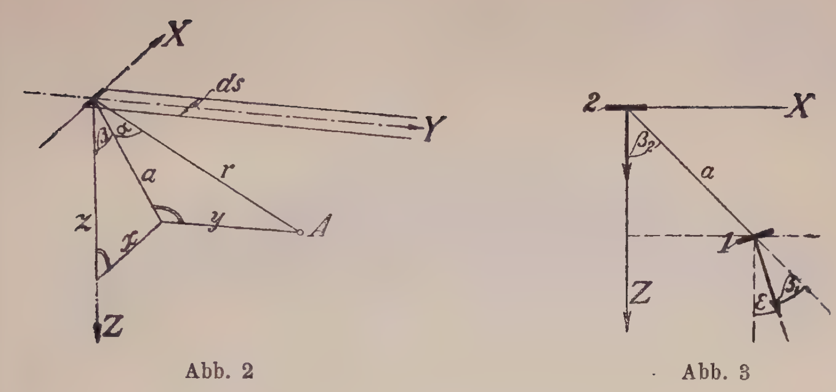


Abb. 2 Abb. 3

$$dw_n = dw \cos\varepsilon + du \sin\varepsilon$$
$$= \frac{\Gamma ds}{4\pi}\left\{(1+\sin\alpha)\frac{\cos(2\beta-\varepsilon)}{a^2} + \frac{\sin a \cos\beta\cos(\beta+\varepsilon)}{r^2}\right\}. \quad (31\,\mathrm{a})$$

Der Widerstand, der an einem tragenden Element 1 durch das Feld des tragenden Elementes 2 entsteht, ist nun

$$d^2 W_{12} = \varrho\,\Gamma_1\, ds_1 \,.\, dw_n\,.$$

Die Ausrechnung ergibt, wenn wieder β_1 für $\beta-\varepsilon$ und β_2 für β eingeführt wird,

$$d^2 W_{12} = \frac{\varrho\,\Gamma_1\, ds_1 \,.\, \Gamma_2\, ds_2}{4\pi}\left\{(1+\sin\alpha)\frac{\cos(\beta_1+\beta_2)}{a^2} + \frac{\sin\alpha\cos\beta_1\cos\beta_2}{r^2}\right\}. \quad (32)$$

Der Widerstand, den das Element 1 an dem Element 2 verursacht, wird, wie leicht ersichtlich, erhalten, wenn für α, β_1 und β_2 $\alpha-\pi$, $\beta_2+\pi$ und $\beta_1+\pi$ gesetzt wird. Hiermit wird

$$d^2 W_{21} = \frac{\varrho\,\Gamma_2\, ds_2 \,.\, \Gamma_1\, ds_1}{4\pi}\left\{(1-\sin\alpha)\frac{\cos(\beta_2+\beta_1)}{a^2} - \frac{\sin\alpha\cos\beta_2\cos\beta_1}{r^2}\right\}. \quad (32\,\mathrm{a})$$

Die beiden Widerstandsanteile sind also n u r d a n n gleich, wenn beide Elemente in derselben Querebene liegen. Sehr bemerkenswert ist, daß die Summe $d^2 W_{12} + d^2 W_{21}$ auch im allgemeinen Fall unabhängig von α ist. Durch Integration über zwei Gruppen von tragenden Elementen ergibt sich demnach die Summe der beiden gegenseitigen Widerstände zu

$$W_{12} + W_{21} = \frac{\varrho}{2\pi}\int\!\!\int \frac{\Gamma_1\,\Gamma_2\, ds_1\, ds_2 \,.\, \cos(\beta_1+\beta_2)}{a^2}, \quad (33)$$

also unabhängig von α d. h. von den y-Werten der einzelnen tragenden Elemente. Diese Summe bleibt daher ungeändert, wenn die beiden tragenden Gruppen in der Flugrichtung gegeneinander

verschoben werden. Da der Gesamtwiderstand eines tragenden Systems außer solchen gegenseitigen Widerständen noch die Eigenwiderstände enthält, die von einer Verschiebung des zugehörigen Flügels in der Flugrichtung natürlich ebenfalls unabhängig sind, läßt sich der Satz aussprechen:

„Der Gesamtwiderstand eines beliebigen Tragwerks[1]) bleibt ungeändert, wenn die tragenden Elemente, ohne ihre Auftriebskräfte zu ändern, in der Flugrichtung gegeneinander schoben werden (oder wenn, wie man sagt, ihre „Staffelung" geändert wird)".

Dieser Satz ist ebenfalls von Munk gefunden und in seiner Göttinger Dissertation[2]) auf andere Art bewiesen worden. Zur Vermeidung von Mißverständnissen sei betont, daß die Anstellwinkel der einzelnen tragenden Elemente bei der Verschiebung entsprechend den Änderungen der normalen Geschwindigkeitskomponenten w_n so zu ändern sind, daß die wirksamen Anstellwinkel (vergl. Nr. 13) und daher auch die Auftriebskräfte unverändert bleiben.

Der innere Grund für die Unabhängigkeit des Gesamtwiderstandes von der Staffelung ist der, daß die Widerstandsarbeit gleich der in der Wirbelbewegung hinter dem Tragwerk zurückgelassenen kinetischen Energie ist; es kommt also nur auf dieses Wirbelsystem selbst an, nicht auf die genaueren Umstände, unter denen es erzeugt worden ist.

Die tragenden Elemente sind bisher immer senkrecht zur Flugrichtung angenommen worden. Eine tragende Linie, die schräg zur Flugrichtung verläuft, kann für die Berechnung von gegenseitigen Widerständen, wenn der Wert $a = 0$ ausgeschlossen bleibt, unbedenklich durch eine aus Elementen der hier betrachteten Art zusammengesetzte Treppenlinie ersetzt werden. Für die Berechnung des Eigenwiderstandes von solchen Flügeln bestehen die in Fußnote S. 17 der I. Mitteilung bereits erwähnten Schwierigkeiten, die eine Durchführung der Integration in endlichen Ausdrücken nur zulassen, wenn von dem Bilde der tragenden Linie abgegangen und und die Verteilung des Auftriebes nach der Tiefe berücksichtigt wird. Nach der vorangegangenen Überlegung über die Wirbel-

1) Mit dem Wort „Tragwerk" werde im Folgenden die allgemeinste Form eines tragenden Flügelsystems bezeichnet, im Gegensatz zu den speziellen Formen des Eindeckers, Zweideckers usw.

2) Isoperimetrische Aufgaben aus der Theorie des Fluges, Göttingen, 1918; z. Z. im Druck.

energie ist es jedoch ersichtlich, daß nicht nur die in diesem Abschnitt abgeleiteten Sätze auch für derartige Flügel gelten, sondern auch deren Eigenwiderstand mit dem der ihnen entsprechenden ganz in der Querebene liegenden tragenden Linie übereinstimmen.

3. Für ein Flügelsystem, bei dem die relative Verteilung des Auftriebs innerhalb jedes einzelnen Flügels festgelegt ist, ergibt sich durch Anwendung von Gl. (16) und (33) der Gesamtwiderstand als eine (definite) quadratische Form der Flügelauftriebe $A_1, A_2 \ldots A_n$. Man kann daher nach Ermittlung der Koeffizienten der Form leicht die Aufgabe lösen, bei gegebenen Gesamtauftrieb $\sum A$ A_1 bis A_n so zu bestimmen, daß der Gesamtwiderstand ein Minimum wird. Die Koeffizienten für Tragwerke, die aus geraden Eindeckern bestehen, deren Mittelpunkte sämtlich in einer die Flugrichtung enthaltenden Lotebene liegen, habe ich unter der Annahme berechnen lassen, daß der Auftrieb über jedem einzelnen Flügel nach einer halben Ellipse verteilt ist, und habe die praktischen Folgerungen aus diesen Ergebnissen in einem Aufsatz in den „Technischen Berichten der Flugzeugmeisterei“ gezogen[1]. Hier mag aus den Ergebnissen das Folgende kurz erwähnt werden: Der gegenseitige Widerstand von zwei Flügeln mit den Spannweiten b_1 und b_2 und dem senkrecht zur Flugrichtung gemessenen Abstand h kann geschrieben werden

$$W_{12} = W_{21} = \frac{\sigma A_1 A_2}{\pi q b_1 b_2}, \tag{34}$$

wobei der Zahlenfaktor σ eine Funktion von $h / \frac{b_1 + b_2}{2}$ und $\mu = b_2/b_1$ ist. Die Ausrechnung ergab folgende Tabelle:

Tabelle 1. Werte von σ.

$2h/(b_1+b_2) =$	0	0,05	0,1	0,15	0,2	0,3	0,4	0,5
$b_2/b_1 = 1{,}0$	1,000	0,780	0,655	0,561	0,485	0,370	0,290	0,230
0,8	0,800	0,690	0,600	0,523	0,459	0,355	0,282	0,225
0,6	0,600	0,540	0,485	0,437	0,394	0,315	0,255	0,210

Für den wichtigsten Fall, $b_1 = b_2$, kann folgende Näherungsformel für σ angewendet werden:

1) „Der induzierte Widerstand von Mehrdeckern“. T. B. Bd. III, S. 309.

$$\sigma = \frac{1 - 0{,}66\,h/b}{1{,}055 + 3{,}7\,h/b}. \tag{35}$$

Als Beispiel mögen die Ergebnisse für einen Doppeldecker mit den Flügelspannweiten b_1 und b_2 ($\leq b_1$) und der Höhe h näher angegeben werden. Der induzierte Widerstand wird

$$W = \frac{1}{\pi q b_1^2}(A_1^2 + 2\sigma\mu A_1 A_2 + \mu^2 A_2^2) \tag{36}$$

dieser wird ein Minimum für

$$A_2 : A_1 = (\mu - \sigma) : \left(\frac{1}{\mu} - \sigma\right);$$

der Wert dieses Minimums ergibt sich zu

$$W_{\text{min.}} = \frac{(A_1 + A_2)^2}{\pi q b_1^2} \cdot \frac{1 - \sigma^2}{1 - 2\sigma\mu + \mu^2}. \tag{37}$$

Der erste Faktor dieser Formel ist der Widerstand eines Eindeckers von der Spannweite b_1 mit dem Auftrieb $A_1 + A_2$. Da für $h > 0$ immer $\sigma < \mu$ ist, wird der zweite Faktor immer kleiner als 1.

Wird die größere Spannweite b_1 vorgeschrieben, die kleinere aber freigelassen, so ergibt sich der kleinste Widerstand für $b_2 = b_1$ und $A_2 = A_1$. Der zweite Faktor von Gl. (37) wird hier $\frac{1}{2}(1 + \sigma)$.

Das Verhältnis des induzierten Widerstandes eines Tragwerks zu dem des Eindeckers mit gleicher Spannweite und gleichem Gesamtauftrieb werde allgemein mit $\varkappa$ bezeichnet. Dann besteht für ein Tragwerk, sobald der zugehörige Zahlwert von $\varkappa$ bekannt ist, die Möglichkeit der Umrechnung von einem Eindecker, für den Versuchsergebnisse vorliegen mögen, in derselben Weise, wie dies in Nr. 15 der I. Mitteilung für die Umrechnung von Eindeckern untereinander dargetan ist. Gleiches gilt auch von Flügelsystemen untereinander. Für den Widerstand von zwei Tragwerken gilt also bei gleichem c_a entsprechend Gl. (27 a):

$$c_{w_1} - c_{w_2} = \frac{c_a^2}{\pi}\left(\frac{\varkappa_1 F_1}{b_1^2} - \frac{\varkappa_2 F_2}{b_2^2}\right) \tag{38}$$ [1].

Dabei ist vorausgesetzt, daß entweder, was in der Regel zutrifft, die Widerstandsziffer c_{w_0} des Profilwiderstandes sich mit dem An-

1) Dieser Gedanke stammt von Munk, vergl. dessen Abhandlung: „Beitrag zur Aerodynamik der Flugzeugtragorgane, Techn. Berichte der Flugzeugmeisterei Bd. II, S. 187.

stellwinkel nicht stark ändert, oder aber, daß die Flächenbelastung aller Flügel dieselbe ist.

4. Bei der Aufgabe, für eine gegebene Auftriebsverteilung eines Mehrdeckers die zugehörige Flügelform zu berechnen, tritt bei den üblichen Maßverhältnissen (Abstand zweier Flügel von der Größenordnung der Flügeltiefe) ein Einfluß sehr fühlbar auf, der bisher noch nicht erwähnt worden ist, der „Krümmungseinfluß“: Die Stromlinien erfahren am Ort eines Flügels durch die Einwirkung der anderen Flügel eine Krümmung. Man darf auf Grund des Superpositionsprinzips, das innerhalb der Theorie der ersten Ordnung als gültig erkannt ist, annehmen, daß ein in der gekrümmten Strömung stehender Flügel dann dieselben Kräfte liefert, wie ein in gerader Strömung stehender, wenn seine eigene Wölbung um den Betrag der Stromlinienkrümmung größer ist als die des Vergleichsflügels. Eine erste Näherung ergibt sich, wenn die Krümmung der Stromlinien für die Mitte der Tiefe unter der Annahme berechnet wird, daß die übrigen Tragflügel durch tragende Linien ersetzt werden dürfen. An einem Aufpunkt A, an dem die zum Flügel und zur Flugrichtung senkrechte Geschwindigkeitskomponente, die von den übrigen Flügeln herrührt, $= w_n$ ist, ist die Krümmung der Stromlinie genähert $\frac{1}{R} = \frac{1}{V}\frac{\partial w_n}{\partial y}$. In dieser Weise ist der Krümmungseinfluß in der Arbeit von Betz über die gegenseitige Beeinflussung der Tragflügel eines Doppeldeckers [1]) berücksichtigt worden. Genauere Werte wird man erhalten, wenn einerseits bei den beeinflussenden Flügeln auf die Verteilung des Auftriebes nach der Tiefe Rücksicht genommen wird, andererseits auch bei dem beeinflußten Flügel statt der Krümmung an einer Stelle in der Mitte der Tiefe der Verlauf von w_n über die ganze Tiefe passend bewertet wird. Derartige Untersuchungen sind zur Zeit in Vorbereitung.

B. Tragwerke kleinsten Widerstandes.

5. Die Bedingungen, unter denen für ein Tragwerk von gegebener Form der Flügelmittellinien ein gegebener Gesamtauftrieb mit einem Minimum von induziertem Widerstand verknüpft ist, hat mein Mitarbeiter M. Munk in seiner bereits erwähnten verdienstvollen Dissertation in voller Allgemeinheit angegeben. Die

1) vergl. Fußnote 4. S. 1 der I. Mitteilung.

im Folgenden mitgeteilte Ableitung, die unter Benutzung der Ergebnisse der Abschnitte 1 und 2 auf kürzestem Wege zum Ziel kommt, unterscheidet sich allerdings formal stark von der Munkschen Ableitung, die auf der Anwendung der Mittel der klassischen Variationsrechnung beruht.

Für das uns vorgegebene Tragwerk dürfen wir, da der Gesamtwiderstand dadurch nicht geändert wird, das durch Verschiebung aller tragenden Elemente in eine Querebene entstandene „ungestaffelte Tragwerk" setzen, für das die Gegenseitigkeitsbeziehung $W_{12} = W_{21}$ besteht. Wir wenden diese Beziehung in der Art an, daß das Tragwerk als Ganzes die eine Gruppe von tragenden Elementen bildet, während die andere Gruppe die ihm überlagerte Variation ist. Diese Einteilung ist zulässig, da der Satz $W_{12} = W_{21}$ nach der Bemerkung am Ende von Anmerkung 3) zu Nr. 1 auch für räumlich zusammenfallende Tragwerke gilt.

Wird der Auftriebsverteilung eine über eine kurze Strecke ds verteilte zusätzliche Luftkraft δN normal zu der gegebenen Richtung des Flügelelements überlagert, und ist die von dem Tragwerk herrührende Störungsgeschwindigkeit normal zum Flügelelement $= w_n$, dann ist der Widerstand W_{21}, den die Geschwindigkeit w_n an der überlagerten Luftkraft hervorbringt, $= \delta N . w_n / V$, der gesamte gegenseitige Widerstand also doppelt so groß.

Zur Erfüllung der Nebenbedingung, daß die Variation des Auftriebes gleich Null ist, seien zwei Zusatzkräfte δN_1 und δN_2 auf zwei tragenden Elementen ds_1 und ds_2 angenommen, die mit der X-Achse die Winkel ε_1 und ε_2 bilden. Dann ist wegen des Verschwindens der ersten Variation des Widerstandes

$$\delta N_1 . w_{n_1} + \delta N_2 . w_{n_2} = 0;$$

die Nebenbedingung ergibt

$$\delta N_1 \cos \varepsilon_1 + \delta N_2 \cos \varepsilon_2 = 0.$$

Beides ist nur verträglich, wenn

$$w_{n_1} : w_{n_2} = \cos \varepsilon_1 : \cos \varepsilon_2$$

oder allgemein für alle tragenden Elemente des ungestaffelten Tragwerks

$$w_n = w_0 \cos \varepsilon \qquad (39)$$

ist. Sind alle tragenden Elemente parallel, dann ist $\cos \varepsilon = 1$ und daher $w_n = w =$ const. für alle tragenden Stellen[1]).

---

1) Für den geraden Eindecker war dieses Ergebnis durch eine Sonderbetrachtung bereits früher (in Nr. 15 der I. Mitteilung) gewonnen worden.

Im allgemeinen Fall von beliebig im Raum verteilten tragenden Elementen führt nach dem Früheren diejenige Auftriebsverteilung zum Widerstandsminimum, die an dem entsprechenden ungestaffelten Tragwerk der Bedingung (36) genügt. Daß es sich hier immer um ein echtes Minimum handelt, ist daraus zu entnehmen, daß die „zweite Variation" den Eigenwiderstand der zusätzlichen Luftkräfte darstellt, also wesentlich positiv ist.

6. Für die Lösung der Aufgabe, die Auftriebsverteilung zu finden, die zum Widerstandsminimum führt, ergibt sich nach Munk ein Weg durch die folgende Betrachtung: Die Geschwindigkeit, die von einem abgehenden Wirbel in der Querebene durch seinen Anfangspunkt hervorgerufen wird, ist gerade die Hälfte derjenigen, die ein nach beiden Seiten unendlich langer Wirbelfaden von gleicher Stärke hervorbringen würde. Bei einem ungestaffelten Tragwerk stimmt daher das Feld der u und w in der Querebene bis auf einen Faktor 2 überein mit demjenigen eines Systems von nach beiden Seiten unendlich langen Wirbeln, das in Anordnung und Stärke dem der abgehenden Wirbel des Tragwerks entspricht. Dieses Feld ist aber eine ebene Strömung, die quellenfrei und außerhalb der durch die tragenden Stellen gezogenen Parallelen zur Flugrichtung auch wirbelfrei ist. Die in der vorigen Nummer abgeleitete Bedingung für das Eintreten des Widerstandsminimums bedeutet für diese ebene Strömung nichts anderes, als die Grenzbedingung für eine aus den Wirbelflächen gebildete starre Zylinderschale, die sich mit der Geschwindigkeit $2\,w_0$ in der Richtung des gegebenen Auftriebs bewegt[1]; die Flüssigkeit ist dabei im Unendlichen in Ruhe.

Die Behandlung von Beispielen ist dadurch zurückgeführt auf eine bekannte Aufgabe des ebenen Problems der Hydrodynamik. Die Wirbelstärke in der Wirbelfläche (d. h. der Geschwindigkeitssprung) für die der Geschwindigkeit $2\,w_0$ entsprechende Bewegung ist nach Früherem $= \frac{d\Gamma}{ds}$; hieraus ergibt sich durch Integration Γ (das an der Flügelspitze $= 0$ ist), gleich dem Potentialsprung $[\Phi]$[2]. Die Normalkraft auf die Längeneinheit des Flügels N ist $\varrho \Gamma V$; somit ist der Auftrieb

$$A = \int N \cos \varepsilon\, ds = \int N\, dx = \varrho V \int [\Phi]\, dx; \qquad (40)$$

---

1) Entsprechend den früheren Festsetzungen werden die Flügelkräfte hier immer in der Richtung gerechnet, in der sie auf die Flüssigkeit wirken.

2) Eckige Klammern als Differenzzeichen!

eine etwaige Seitenkraft ist

$$S = \int N \sin \varepsilon \, ds = -\int N dz = -\varrho V \int [\Phi] \, dz\,^{1}). \qquad (40\,\text{a})$$

Man kann die Strömung in eine stationäre verwandeln, indem man der Flüssigkeit eine der Geschwindigkeit des starren Gebildes entgegengesetzt gleiche Geschwindigkeit erteilt, durch die das starre Gebilde zur Ruhe gebracht wird, und im Unendlichen die Geschwindigkeit $-2w_0$ entsteht. Die Gleichung (40) wird hierdurch nicht geändert. —

Für mehrfach zusammenhängende tragende Gebilde erhält man bei diesem letzteren Bilde in dem abgeschlossenen Inneren ruhende Flüssigkeit. Der Potentialsprung wird dadurch, daß für das Potential im Innern ein beliebiger konstanter Wert angenommen werden kann, unbestimmt. Dem entspricht in unserer Anwendung die Tatsache, daß bei einem ringförmig geschlossenen Tragwerk eine konstante Normalkraft N, die ja auch keine von Null verschiedene Resultante ergibt, keinen Widerstand hervorruft und daher auch zu einer bestehenden Verteilung von Auftrieb willkürlich hinzugefügt werden darf, ohne den Widerstand zu ändern. Man kann auch das Innere des Ringes durch tragende Linien, die von einem Punkt der Berandung zu einem anderen führen, in Felder teilen, und in jedem Feld ein anderes Potential annehmen. Die durch die jetzt vorhandenen Potentialsprünge definierten Normalkräfte ändern den Widerstand ebenfalls nicht. Der innere Grund dieses Verhaltens ist der, daß von den inneren Flügeln, bei denen auf der ganzen Länge die Normalkraft, also auch Γ konstant ist, keine Wirbel abgehen.

Für die praktische Rechnung wird man die Geschwindigkeit im Unendlichen (nach der zweiten Auffassung), die $= 2w_0$ ist, auf 1 normieren, d. h. man berechnet statt Φ das reduzierte Potential $\varphi = \Phi/2w_0$, das die Dimension einer Länge hat. $\int [\varphi]\, dx$ ist also eine Fläche, die im Folgenden F' heißen möge. Hiermit wird

$$A = \varrho V \,.\, 2w_0 F',$$

also

$$w_0 = A/2\varrho V F'; \qquad (41)$$

1) Setzt man $dx + i\,dz = d\xi$, $\Phi + i\Psi = Z$ und beachtet, daß der Sprung der Stromfunktion Ψ aus Gründen der Kontinuität $= 0$ ist, so kann man auch schreiben $A - iS = \varrho V \oint Z\, d\xi$, wobei der Integrationsweg das starre Gebilde — oder anders gesprochen, das Tragwerk — von außen umschlingt. Für das Integral kann auch das Residuum im Unendlichen genommen werden.

andererseits ist

$$W = \int \frac{w_n}{V} N ds = \frac{w_0}{V} \int \cos \varepsilon \,.\, N ds = \frac{w_0}{V} A$$

(gemäß Gl. 39 und 40); also mit (41)

$$W = A^2/2\varrho V^2 F' = A^2/4qF'. \tag{42}$$

Werden wieder die dimensionslosen Größen $c_a = A/qF$ und $c_{w_i} = W/qF$ eingeführt, so wird

$$c_{w_i} = c_a^2 F/4F'. \tag{43}$$

7. Das Tragwerk erteilt, während es seine Bahn entlang bewegt wird, einer Luftpartie nach der anderen die, wie wir annehmen, kleinen Geschwindigkeiten, die wir als das Ergebnis der abgehenden Wirbel kennen gelernt haben. Gemäß der vorangegangenen Überlegung — oder auch gemäß den Gleichungen (31) mit $\sin \alpha = 1$ — sind diese hinreichend weit hinter dem Tragwerk, wenn wir von den bis dahin eingetretenen Deformationen der Wirbelfläche im Sinne der Theorie der I. Ordnung absehen, $= 2u$ bezw. $2w$, wobei gemäß unserer bisherigen Bezeichnung u und w die Geschwindigkeiten beim ungestaffelten Tragwerk in der Querebene sind. Statt diese Geschwindigkeiten in zeitlicher Reihenfolge immer neuen Luftpartien zu erteilen, kann man sich vorstellen, daß sie durch einen Stoß auf der ganzen Bahn des Tragwerks gleichzeitig erzeugt werden. Zur Ausführung dieses Stoßes bedarf man eines festen zylindrischen Gebildes von der Gestalt des vom Tragwerk überstrichenen geometrischen Ortes. Für den Fall, daß es sich um ein Tragwerk kleinsten Widerstandes handelt, ist dieses Gebilde starr, andernfalls wird es bei dem Stoß eine Formänderungsbewegung ausführen. Man kann nun leicht von dem Stoßvorgang auf den Vorgang am Tragwerk umrechnen und wird dabei wieder auf Gl. (40) geführt. Die Stoßdrücke $\int_0^\tau p\, d\tau$ (p = Flüssigkeitsdruck, τ = Stoßdauer), die nach einem bekannten Satz der Hydrodynamik $= \varrho \Phi$ sind, wo Φ das Strömungspotential ist, ergeben nämlich durch den Drucksprung $[p]$ an den festen Flächen eine Stroßkraft normal zur Fläche, die für die Flächeneinheit $= \int_0^\tau [p]\, dt = \varrho [\Phi]$ ist. Diese entspricht der Wirkung der Normalkraft N für die Zeit, in der das Tragwerk die Längen-

einheit zurücklegt, es ist somit $N/V = \varrho[\Phi]$ in Übereinstimmung mit dem Früheren.

Für die Fläche F' läßt sich noch eine einfache mechanische Bedeutung angeben. Schreibt man die Geschwindigkeit des zylindrischen Gebildes nach dem Stoß, also die Geschwindigkeit der Wirbelfläche weit hinter dem Tragwerk, $2w_0 = w_1$, dann ist nach Gl. (41) und (42)

$$A = \varrho F' V w_1$$

und

$$W.V = A w_0 = \varrho F' V \frac{w_1^2}{2}.$$

$\varrho F' V$ ist die sekundlich durch den Querschnitt F' hindurchtretende Luftmasse. Die Gleichungen zeigen, daß man den Auftrieb und die Widerstandsarbeit richtig erhält, wenn man annimmt, daß die Luft innerhalb des Querschnittes F' die Abwärtsgeschwindigkeit w_1 erhalten hat, außerhalb aber unabgelenkt bleibt.

8. Der vergleichenden Beurteilung von verschiedenen Tragwerkformen dient die nachfolgende Betrachtung:

Zunächst gilt der Satz: „Bei unverändertem Gesamtauftrieb wird durch Hinzufügen von Tragwerkteilen das Minimum des induzierten Widerstandes verkleinert, durch Wegnahme von Tragwerkteilen dagegen vergrößert. In Ausnahmefällen kann es unverändert bleiben". Der Beweis ergibt sich durch folgende Überlegung: das Weglassen eines Tragwerkteiles ist gleichbedeutend damit, daß für das vollständige Tragwerk eine Auftriebsverteilung gewählt wird, bei der die Auftriebsdichte in dem fortzulassenden Teil Null ist. Diese Verteilung wird im allgemeinen von der Minimalverteilung des vollständigen Tragwerks abweichen, ihr Widerstand also größer sein, als der Minimalwiderstand des vollständigen Tragwerks. Nur in dem Ausnahmefall, daß eine zum Widerstandsminimum führende Auftriebsverteilung mit Auftrieb Null in den wegzulassenden Teilen möglich ist, bleibt der Widerstand unverändert.

Nach dem in Nr. 6 über mehrfach zusammenhängende Tragwerke Gesagten dürfen im Innern von ringförmigen Tragwerken Flügel, die von einem Randpunkt zu einem anderen verlaufen, nach Belieben eingefügt werden, ohne daß der Minimalwiderstand geändert wird. Durch solche Zufügungen und nachheriges Fortlassen des Überflüssigen kann man aber aus einem ringförmigen Tragwerk alle beliebigen Tragwerke herstellen, die ganz in seinem Innern liegen, oder wenigstens seine Grenzen nirgends überschreiten. Der induzierte Widerstand kann bei dieser Umwandlung nie abnehmen,

im allgemeinen wird er beim Fortlassen der in Wegfall kommenden Teile zunehmen. Hieraus entnimmt man den für Abschätzungen wertvollen Satz: „Ein ringförmiges Tragwerk besitzt von allen Tragwerken, die seine Grenzen nirgends überschreiten, den kleinsten Minimalwert des induzierten Widerstandes".

9. Die Fläche F' läßt sich nach den aus der Hydrodynamik bekannten Methoden für eine Reihe von Tragwerksformen berechnen. Für den geraden Eindecker von der Spannweite b, ententsprechend der Strömung um eine ebene Platte, ergibt sich F' gleich der Kreisfläche mit b als Durchmesser: $F' = \frac{\pi}{4} b^2$. Durch Einsetzen dieses Wertes in Gl. (42) ergibt sich die früher gewonnene Beziehung (25). Weiter konnte F' für ein Tragwerk mit einem lotrechten Kreis als Mittellinie leicht berechnet werden. Ist der Durchmesser $= b$, so wird hier $F' = \frac{\pi}{2} b^2$.

Die Ermittlung von F' für einige verwickeltere Fälle, bei denen elliptische Integrale auftreten, hat Herr R. Grammel auf meine Veranlassung durchgeführt. Im Folgenden mögen die Ergebnisse ohne Beweis kurz angegeben werden. Dabei bedeuten K und E die Normalintegrale erster und zweiter Gattung zum Modul k, K und E die zugehörigen vollständigen Normalintegrale; K' E', K' und E' sind die entsprechenden Integrale zum Modul $k' = \sqrt{1-k^2}$.

1) Für einen aus zwei geraden Eindeckern von gleicher Spannweite zusammengesetzten Doppeldecker ergibt sich folgende Rechenvorschrift: für einen Modul k, der nahe an Null angenommen wird, ermittle man die Hilfsgrößen

$$\varepsilon^2 = \mathsf{E}'/\mathsf{K}'; \quad t = \sqrt{\frac{1-\varepsilon^2}{1-k^2}} \text{ und } 1/\delta = E'(t) - \varepsilon^2 K'(t);$$

dann ergibt sich das Verhältnis der Höhe zur Spannweite des Doppeldeckers zu

$$h/b = \delta\pi/2\,\mathsf{K}'$$

und die Fläche F' zu

$$F' = \frac{\pi}{4} b^2 \delta^2 (1 + k^2 - 2\varepsilon^2).$$

2) Für ein rechteckförmiges ringartig geschlossenes Tragwerk (das also aus einem Doppeldecker durch Einfügen von abschließenden Seitenwänden entsteht, und das nach dem in Nr. 8 Gesagten einen kleineren induzierten Widerstand hat als alle in seinem Rahmen eingeschlossenen Mehrdecker) fand Grammel die Beziehungen

$$h/b = \frac{\mathsf{E} - k'^2 \mathsf{K}}{\mathsf{E}' - k^2 \mathsf{K}'} \text{ und}$$

$$F' = \frac{\pi b^2 k'^2}{4(\mathsf{E}' - k^2 \mathsf{K}')^2}.$$

3. Für einen Eindecker von der Spannweite b mit einem Spalt von der Breite d in der Mitte wird zunächst zu der vom Mittelpunkt aus gerechneten Abszisse x die Hilfsvariable

$$t = \sqrt{\frac{b^2 - 4x^2}{b^2 - d^2}} \text{ gebildet, dann mit } k = \sqrt{1 - (d/b)^2}$$

$$\Phi = \frac{b}{2}\left(\frac{\mathsf{E}}{\mathsf{K}} K(t) - E(t)\right)$$

berechnet, worauf F' durch Quadratur aus

$$F' = 4 \int_{d/2}^{b/2} \Phi \, dx$$

ermittelt werden kann.

Die Durchführung der Zahlenrechnung für diese drei Fälle, die ich Herrn K. Pohlhausen verdanke, hat die folgenden nicht uninteressanten Zahlenwerte ergeben. Es ist überall das Verhältnis $\varkappa$ des Widerstandes der untersuchten Anordnung zu dem Widerstand des Eindeckers von gleicher Spannweite angegeben:

Tabelle 2. Werte von $\varkappa = \frac{\pi}{4} b^2 : F'$.

$h/b = 0$		0,05	0,10	0,15	0,2	0,3	0,4	0,5
Doppeldecker	1,000	0,890	0,825	0,776	0,739	0,684	0,645	0,615
Rechteck	1,000	0,865	0,787	0,728	0,680	0,601	0,545	0,500

Der Doppeldecker ist also bei gleicher Spannweite und gleichem Auftrieb dem Eindecker, das Rechteck beiden überlegen, wie das allgemeine Gesetz es verlangt. Für den praktischen Gebrauch dienen folgende Näherungsformeln: für den Doppeldecker:

$$\varkappa \cong \frac{1 + 1{,}63\, h/b}{1{,}027 + 3{,}84\, h/b} \tag{44}$$

für das Rechteck

$$\varkappa \cong \frac{1 + 0{,}45\,h/b}{1{,}045 + 2{,}8\,h/b}. \tag{45}$$

Bei dem Tragflügel mit Spalt liegt der Vergleich mit demjenigen Eindecker nahe, der durch Schließen des Spaltes unter Zusammenschieben der beiden Teile entsteht, der also die Spannweite $b-d$ hat. Über das Verhältnis des Widerstandes mit Spalt zu dem ohne Spalt gibt die Tabelle 3 Auskunft.

Tabelle 3. Werte von $\varkappa = \frac{\pi}{4}(b-d)^2 : F'$.

d/b = 0,000	0,001	0,01	0,0316	0,1	0,25	0,5	1,0
$\varkappa$ = 1,000	1,312	1,480	1,612	1,763	1,896	1,975	2,000

Diese Werte werden durch die Formel angenähert:

$$1/\varkappa = 1 - \frac{1}{2\sqrt{1 + 0{,}35\,({}^{10}\log b/d)^2}}. \tag{46}$$

C. Einfluss von Wänden und von freien Grenzen.

10. Das Studium des Einflusses einer seitlichen Begrenzung des von der Flüssigkeit erfüllten Raumes auf einen darin befindlichen Tragflügel ist vor allem von Bedeutung für die Beurteilung von Versuchsergebnissen, die im künstlichen Luftstrom gewonnen sind. Die Versuchsanstalten verwenden teils Kanäle mit festen Wänden, teils freie Luftstrahlen, die mit angenähert zylindrischer Begrenzung den sie umgebenden Luftraum durchströmen. Es entsteht so die Aufgabe, das Verhalten eines Tragflügels oder Tragwerks bei festen oder freien Grenzen parallel der Richtung von V zu untersuchen.

Bei festen Wänden ist die exakte Bedingung das Verschwinden der Normalkomponente w_n der Geschwindigkeit an den festen Wänden, bei freien Grenzen dagegen die Unveränderlichkeit des Druckes auf diesen. Die letztere Bedingung kann im Bereich der Theorie der I. Ordnung wie folgt umgeformt werden: Nach dem Bernoulli'schen Satz ist an den freien Grenzen

$$(V+v)^2 + u^2 + w^2 = \text{const.} = V^2$$

oder

$$2Vv + u^2 + v^2 + w^2 = 0;$$

innerhalb der I. Ordnung gilt also $v = 0$ auf den freien Grenzen, oder auch $v = 0$ auf demjenigen Zylinder parallel zur Y-Achse [1]), der bei Abwesenheit der Störung durch den Tragflügel die Strahlgrenze darstellen würde.

Die beiden Aufgaben lassen sich in der Weise lösen, daß zunächst der Tragflügel mit dem aus den früheren Abschnitten bekannten Geschwindigkeitsfeld für die unendlich ausgedehnte Flüssigkeit betrachtet wird, und nun (im Innern singularitätenfreie) Potentialströmungen zugefügt werden, bei denen an den Grenzen w_n bezw. v entgegengesetzt gleiche Werte haben, wie die Flügelströmung. Die Geschwindigkeitskomponenten, die von der Zusatzströmung am Flügel verursacht werden, ändern dort den Anstellwinkel und den Widerstand in gleicher Weise, wie dies früher von den Störungsgeschwindigkeiten dargelegt worden ist, die von andern Flügeln herstammen.

Für den freien Strahl, der uns wegen der Göttinger Versuchsanlage besonders naheliegt, ergibt sich bei beliebiger Form des Strahlquerschnittes das Folgende: Die Randwerte der Y-Komponente der Flügelströmung auf der zylindrisch angenommenen Strahloberfläche seien $\bar{v}$; die Y-Geschwindigkeit der zusätzlichen Potentialbewegung auf dem Strahlrand, ist daher $\frac{\partial \Phi}{\partial y} = -\bar{v}$; das Potential wird also längs jeder Erzeugenden des Strahlrandes integrierbar: es ist

$$\bar{\Phi}(y) = -\int_{-\infty}^{y} \bar{v}\, dy.$$

(Die untere Grenze des Integrals ergibt sich daraus, daß hinreichend weit stromaufwärts das zusätzliche Feld verschwinden muß). Das zusätzliche Feld im Innern des Strahles ergibt sich hieraus durch Lösung der „ersten Randwertaufgabe“.

Soweit es sich lediglich um die Berechnung des Gesamtwiderstandes handelt, genügt es wieder, den Zustand weit hinter dem Tragwerk, also für $y = \infty$ zu kennen. Die Randwerte

$$\bar{\Phi}_1 = -\int_{-\infty}^{+\infty} \bar{v}\, dy$$

1) Das Koordinatensystem möge entsprechend dem bisher gebrauchten so gelegt sein, daß die Y-Achse mit der Richtnng des ankommenden Luftstrahles übereinstimmt und die Z-Achse parallel zum resultierenden Auftrieb ist.

sind, wie leicht zu sehen, unabhängig von etwaiger Staffelung der tragenden Elemente, aus deren Wirkungen sich $\bar{v}$ additiv zusammensetzt. Diese vereinfachte Aufgabe führt also wieder auf ein ebenes Problem, nämlich auf die Ermittlung des Geschwindigkeitsfeldes $u_1\, w_1$ für die Randwerte $\overline{\Phi}_1$ des Potentials.

Für ein ungestaffeltes Tragwerk ist v, und daher auch $\bar{v}$, eine gerade Funktion von y, damit wird $-\int_0^y \bar{v}\, dy$ eine ungerade Funktion von y. Das zusätzliche Randpotential wird damit

$$\overline{\Phi} = -\int_{-\infty}^{0} \bar{v}\, dy - \int_0^y \bar{v}\, dy = \overline{\Phi}_0 + \overline{\Phi}', \text{ wo } \overline{\Phi}' \text{ ungerade.}$$

Entsprechendes gilt für das Potential im Innern; die Geschwindigkeiten u und w zerlegen sich dadurch in einen von y unabhängigen Teil $u_0 = \frac{\partial \Phi_0}{\partial x}$, $w_0 = \frac{\partial \Phi_0}{\partial z}$ und in einen Teil $u' = \frac{\partial \Phi'}{\partial x}$, $w' = \frac{\partial \Phi'}{\partial z}$, der in y ungerade ist. In der Querebene des Tragwerks verschwinden die ungeraden Teile und es bleibt somit wieder das ebene Problem übrig. Wie leicht zu sehen, ist genau entsprechend dem Früheren $\Phi_1 = 2\,\Phi_0$ und $u_1 = 2\,u_0$, $w_1 = 2\,w_0$.

Wegen $\bar{v} = 0$ am Rande für die fertige Strömung ist dort überall $\Phi = \text{const.}$, also setzen die Stromlinien der zu der Parallelbewegung mit der Geschwindigkeit V hinzukommenden Strömung senkrecht auf der Strahloberfläche auf. Bei der Betrachtungsweise von Nr. 7, die hier ebenfalls anwendbar ist, findet man aus der Forderung, daß der Druck auf der Strahloberfläche während des Stoßvorganges sich nicht ändert, wegen $\Phi = \int_0^\tau p\, dt$ wieder $\Phi = \text{const.}$ für die Strahloberfläche.

Die Bedingungen für den Kanal lassen sich ebenso durchführen. Nur wird man hier wegen $w_n = 0$ auf die „zweite Randwertaufgabe“ geführt. u und w bestehen auch hier aus einer ebenen Strömung plus einem in y ungeraden Anteil, wenn das Tragwerk ungestaffelt ist.

Für die Ermittlung der ebenen Strömung $u_0\, w_0$ läßt sich mit Vorteil die Methode der Spiegelbilder einwenden. Bei einem Strahl oder Kanal von Kreisquerschnitt hat man dabei in bekannter Weise zu jedem abgehenden Wirbel einen ebenso starken von entgegengesetztem (bezw. gleichem) Vorzeichen in dem Punkte hinzuzunehmen, der dem anderen nach reziproken Radien entspricht. Einem Streifen des abgehenden Wirbelbandes von der Stärke

$\frac{d\Gamma}{ds}\,ds$ entspricht somit ein gleich starker vom Betrage $\frac{d\Gamma'}{ds'} \cdot ds'$. Durch Integration zwischen je zwei entsprechenden Punkten ergibt sich hieraus, daß für die gespiegelten tragenden Wirbel $\Gamma' = \mp\Gamma$ ist (Minuszeichen für den Strahl, Pluszeichen für den Kanal).

11. Die zahlenmäßigen Verhältnisse für den Kreisstrahl habe ich für den einfachsten Fall, einen geraden Eindecker in der Mitte des Strahles, näher untersucht, und zwar unter der Annahme, daß der Auftrieb nach einer halben Ellipse verteilt ist. Ist b die Spannweite des Eindeckers und D der Strahldurchmesser, so wird die von den Strahlgrenzen verursachte Störungsgeschwindigkeit w' in der Entfernung x aus der Mitte, mit der Abkürzung $\xi = 2\,x\,b/D^2$

$$w' = \frac{A}{4\pi R^2 \varrho V}\left(1 + \frac{3}{4}\xi^2 + \frac{5}{8}\xi^4 + \frac{35}{128}\xi^6 + \cdots\right) \qquad (48)$$

Hierdurch ergibt sich ein Zusatzwiderstand

$$W' = \frac{A^2}{4\pi R^2 \varrho V^2}\left(1 + \frac{3}{16}\left(\frac{b}{D}\right)^4 + \frac{5}{64}\left(\frac{b}{D}\right)^8 + \cdots\right) \qquad (49)$$

Eine Vergleichsrechnung mit rechteckiger Auftriebsverteilung ergab das erste Glied in der Widerstandsformel identisch mit dem obigen; der Schluß liegt nahe, daß dies auch bei irgend einem anderen Tragwerk der Fall ist. Wenn man sich auf dieses erste Glied beschränkt, was praktisch in den meisten Fällen ausreicht, so läßt sich der gesamte induzierte Widerstand unter Einführung des Querschnitts F_0 des Luftstromes schreiben:

$$W = \frac{A^2}{4q}\left(\frac{1}{F'} + \frac{1}{2F_0}\right)^{1)}. \qquad (50)$$

Bei einem Kanal von Kreisquerschnitt erhält man die gleichen Störungsbeiträge, aber mit entgegengesetztem Vorzeichen. Das erste Glied der Reihe bringt eine recht merkliche Korrektur, bei $b/D = 1/2$ z. B. wird sie bereits 1/8 des induzierten Widerstandes. Der vollständige Ausdruck liefert anstelle von 0,125 bei der elliptischen Verteilung 0,1262 und bei der Rechteckverteilung 0,127.

12. Der durch den Kreisstrahl hindurchgesteckte Flügel hat ebenfalls ein gewisses praktisches Interesse, da er eine besonders

1) Der Faktor 2 bei F_0 läßt sich so deuten, daß der Strahl nur die Hälfte des Auftriebes als Impuls durch Ablenkung von der ursprünglichen Richtung aufnimmt, da die andere Hälfte — einer einfachen Rechnung zufolge — durch den Impuls des zum Flügel gehörigen Wirbelsystems aufgenommen wird.

bequeme Versuchsanordnung ergibt. Man kann seine Theorie erhalten, wenn man beachtet, daß da, wo der Flügel durch die Strahloberfläche hindurchtritt, wegen der Druckgleichheit über und unter dem Flügel der Auftrieb Null werden muß. Für eine Reihenentwicklung ist hier die Funktion $\frac{R^2 - x^2}{R^2 + x^2}$ und ihre ungeraden Potenzen geeignet, da die Reihe für $x = \pm R$ verschwindet und außerdem der Bedingung für den Kreisstrahl $f(R^2/x) = -f(x)$ genügt. Die ziemlich umständliche Durchrechnung für die zwei ersten Glieder, der sich Herr K. Pohlhausen in dankenswerter Weise unterzog, ergibt für

$$\Gamma = VR\left\{c_1 \frac{R^2 - x^2}{R^2 + x^2} + c_2 \left(\frac{R^2 - x^2}{R^2 + x^2}\right)^3\right\}$$

$$A = \varrho V^2 R^2 \left\{c_1(\pi - 2) + c_2\left(\tfrac{3}{4}\pi - 4\right)\right\} \tag{51}$$

$$w = R^2 V \frac{R^2 - x^2}{(R^2 + x^2)^2}\left\{\frac{c_1}{2} + \frac{3}{4} c_2 \cdot \frac{R^4 - 6R^2 x^2 + x^4}{(R^2 + x^2)^2}\right\} \tag{52}$$

$$W = \frac{\pi}{8} \varrho V^2 R^2 \left\{c_1^2 + \tfrac{3}{2} c_1 c_2 + \tfrac{3}{4} c_2^2\right\}. \tag{53}$$

Das Minimum von W bei gegebenem A wird erhalten für

$$c_1 = 1{,}213\, A/\varrho V^2 R^2 \text{ und } c_2 = -0{,}541\, A/\varrho V^2 R^2$$

und wird

$$W_{\text{min.}} = 0{,}277\, A^2/\varrho V^2 R^2 = 1{,}74\, A^2/4q F_0. \tag{54}$$

Dieses ist das 1,74fache des Widerstandes eines Eindeckers von der Spannweite $2R$ in der unendlich ausgedehnten Flüssigkeit (die einfache Näherungsformel (50) gibt das 1,5fache, was für diesen extremen Fall immerhin noch recht annehmbar ist).

13. Bei Tragwerken von solcher Tiefenentwicklung, daß es nicht mehr erlaubt ist, die Veränderlichkeit der Störungsgeschwindigkeit mit der Tiefe zu vernachlässigen, muß der Wölbungseinfluß beachtet werden. Ein Fall, für den dieser ermittelt worden ist, der aber auch sonst praktische Bedeutung hat, mag hier noch kurz behandelt werden. Um das ebene Problem im Versuch angenähert zu verwirklichen, haben wir einen Tragflügel seitlich durch zwei parallele Wände abgeschlossen, die den Strahl in seiner ganzen Höhe auf eine größere Strecke begleiten. Die Ober- und Unterseite des Strahles war dabei frei. Die Bewegung ist eine ebene Strömung mit zwei freien Grenzen über und unter dem Flügel. Dieser befand sich in der Strahlmitte. Durch unendlich

oft wiederholte Spiegelung an der oberen und unteren Grenze ergibt sich hier eine unendliche Reihe senkrecht übereinander befindlicher gleicher und gleichsinniger tragender Wirbel von der Stärke Γ in Abständen gleich der Strahlhöhe h. Die Zusatzgeschwindigkeit im Unendlichen hinter dem Tragflügel wird

$$w_1 = 2w_0 = A/\varrho V l h = A/\varrho V F_0, \tag{55}$$

der Widerstand also

$$W = A \cdot w_0/V = A/2\varrho V^2 F_0 = A/4q F_0. \tag{56}$$

Zur Ermittlung des Wölbungseinflusses werde die Geschwindigkeit w für die Strahlmittellinie ($\parallel V$) berechnet. Es wird durch die Wirkung der gespiegelten Wirbel für $z = 0$

$$w = w_0 = \frac{\Gamma y}{\pi} \sum_1^\infty \frac{1}{n^2 h^2 + y^2}; \tag{57}$$

für kleine y ist demnach

$$w - w_0 \cong \frac{\Gamma y}{\pi h^2} \sum_1^\infty \frac{1}{n^2} = \frac{\pi}{6} \frac{\Gamma y}{h^2},$$

also die Krümmung

$$1/R = \frac{\partial}{\partial y}\left(\frac{w}{V}\right) = \frac{\pi \Gamma}{6 V h^2} = \frac{\pi A}{6 \varrho V^2 l h^2} = \frac{\pi c_a t}{12 h^2}. \tag{58}$$

Um diesen praktisch allerdings meist geringen Betrag[1]) muß die Wölbung des Versuchsflügels größer sein als die des unendlich langen Flügels in unbegrenzter Flüssigkeit, dessen Eigenschaften ermittelt werden sollen.

Beim ebenen Probleme im Kanal erhält man mit unendlich vielen gespiegelten Wirbeln von abwechselnden Vorzeichen $w_1 = 0$, also keinen induzierten Widerstand, ferner entsprechend

$$w = \frac{\Gamma y}{\pi} \sum_1^\infty \frac{-1^n}{n^2 h^2 + y^2} \tag{59}$$

für $z = 0$, was eine Wölbungskorrektur von entgegengesetztem Sinn und vom halben Betrage wie beim Strahl ergibt.

---

1) Für zahlenmäßige Abschätzungen sei bemerkt, daß die flugtechnisch interessierenden Werte von c_a zwischen 0,2 und 1,2 liegen.

D. Die Zustände in grossem Abstand vom Flügel.

14. Für die Ermittlung des Geschwindigkeitsfeldes eines Tragwerks in großer Entfernung von diesem genügt es, wenn die Umgebung des Wirbelschweifes zunächst außer Acht bleiben darf, die Formeln (31) unter der Näherungsannahme zu integrieren, daß die Fahrstrahlen vom Aufpunkt nach den einzelnen tragenden Elementen parallel sind. Sind die Auftriebsrichtungen der einzelnen tragenden Elemente um Winkel ε gegen die Z-Achse geneigt (vgl. Fig. 3), so ist überall anstelle von β zu setzen $\beta - \varepsilon$ und ferner zu beachten, daß die in Gl. (31) angegebenen Beträge nach ihrem tragenden Element orientiert sind. Die Integration ergibt, wenn man das neue Koordinatensystem so wählt, daß die Seitenkraft $S = \int \Gamma ds \sin \varepsilon$ verschwindet, Formeln, die sich von den Gl. (31) nur dadurch unterscheiden, daß $\int \Gamma ds \cos \varepsilon = \int \Gamma dx = A/\varrho V$ anstelle von Γds tritt.

Der Druck in großer Entfernung vom Tragwerk wird — abgesehen von dem Gebiete des Wirbelschweifes — $p = p_0 - \varrho V v$, da die Quadrate der Störungsgeschwindigkeiten hier vernachlässigt werden dürfen, also gemäß Gl. (31)

$$p - p_0 = \frac{\varrho V}{4\pi} \int \Gamma dx \cdot \frac{\cos \alpha \cos \beta}{r^2} = \frac{Az}{4\pi r^2}. \tag{60}$$

Unter dem Tragwerk herrscht Überdruck, über ihm Unterdruck.

Das Integral des Druckunterschieds über eine unendlich ausgedehnte Ebene senkrecht zur Z-Achse (wagerechte Ebene) wird unabhängig vom Abstand der Ebene $= \pm \frac{1}{2} A$, je nachdem die Ebene unter oder über dem Tragwerk liegt. Zwei Ebenen dieser Art, die das Tragwerk zwischen sich schließen, nehmen also zusammen den ganzen Auftrieb in Form von Druck auf. Dies gilt so für ein nach allen Richtungen unendlich ausgedehntes Luftmeer. Ist in endlicher Entfernung ein wagerechter Erdboden vorhanden, dann muß, um der Grenzbedingung dort zu genügen, ein „gespiegeltes Tragwerk“ im Bildpunkt des wirklichen angenommen werden. Dieses erzeugt in jeder wagerechten Ebene über dem Boden einen Überdruck mit einer Resultierenden, die gleichfalls $\frac{1}{2} A$ ist. Mit dem Druckfeld des wirklichen Tragwerks zusammen liefert demnach jede Ebene zwischen dem Tragwerk und dem Erdboden, also auch dieser selbst, den Betrag A, jede Ebene über dem Tragwerk dagegen, wo die Überdruckwirkung des gespiegelten Tragwerks die Unterdruckwirkung des wirklichen im Gesamtbetrag aufhebt, den Betrag Null.

15. Von den Zuständen im Wirbelschweif, die mit dem Widerstand in Beziehung stehen, soll später die Rede sein. Was den Auftrieb betrifft, so interessiert noch die Rolle des in der Flüssigkeit vorhandenen Impulses. Dieser ist durchaus nicht eindeutig bestimmbar, sondern, wie schon das Verhalten der Druckresultierenden andeutete, von den Verhältnissen im Unendlichen abhängig. Umgibt man das in einem allseitig unbegrenzten Medium befindliche Tragwerk mit einer parallelepipedischen Kontrollfläche, so ergibt die Anwendung des Impulssatzes für stationäre Bewegung[1]), wenn man zuerst die Grenzflächen senkrecht zur Y- und Z-Achse und dann erst die senkrecht zur X-Achse ins Unendliche rückt, einen Impuls gleich dem Auftrieb, der aus dem Wirbelschweif stammt; geht man zuerst in der X- und Z-Achse, dann erst in der Y-Achse ins Unendliche, so liefert der Wirbelschweif nichts, dafür aber der tragende Wirbel einen Impuls $= A$; wenn endlich zuerst in der X- und Y-Achse ins Unendliche gegangen wird, und erst alsdann in der Z-Achse, so wird der Impuls Null und es ergeben sich nur Druckkräfte mit der Resultante A. In anderen Fällen wird Druck und Impuls zusammen erhalten. Von der Art des Grenzüberganges unabhängig wird das Ergebnis im Falle eines Bodens in endlicher Entfernung und in dem Fall, daß der Flug in einer allseitig unendlich ausgedehnten Flüssigkeit erst eine endliche Wegstrecke hindurch andauert. Im ersteren Fall ergibt sich ein Bodendruck $= A$ und kein Impuls, im zweiten Fall nur Impuls im Betrage A.

Dieses letztere Ergebnis ist am besten mittelst eines bekannten Satzes über die Bewegungsgröße, die zu einer geschlossenen Wirbellinie von der Stärke Γ gehört, einzusehen. Diese Bewegungsgröße ist, wenn $\mathbf{F}$ die als Vektor genommene von der Wirbellinie berandete Fläche ist, für eine nach allen Seiten unendlich ausgedehnte Flüssigkeit $= \varrho \Gamma \mathbf{F}$. Wenn der Flug mitten in einer allseitig unbegrenzten Flüssigkeit begonnen worden ist, so ist der

1) Vgl. etwa meinen Artikel „Flüssigkeitsbewegung“ im „Handwörterbuch der Naturwissenschaften“, S. 112. — Die dort (S. 114) gemachte Bemerkung, daß der Auftrieb von der Luft zunächst in Form von Impuls aufgenommen und erst später, wenn die abwärts bewegten Luftmassen den Boden erreichen, in Form von Druck auf diesen übertragen wird, war irrig. Der Irrtum hing mit der oben dargelegten Mehrdeutigkeit des Impulses zusammen, die mir damals noch nicht bekannt war. — Analoge Betrachtungen für die Hubschraube zeigen übrigens, daß auch hier der Auftrieb momentan — oder wenn die Kompressibilität berücksichtigt wird, mit Schallgeschwindigkeit — als Druck auf den Boden übertragen wird.

Wirbelschweif jedes Flügels hinten durch einen „Anfahrwirbel", vorne aber durch den „tragenden Wirbel" geschlossen. Dies gilt auch von jeder einzelnen Wirbellinie von der Stärke $d\Gamma$. Ist l der Abstand der beiden zusammengehörigen Wirbellinien — oder falls dieser Schwankungen unterworfen ist, der Mittelwert dieses Abstandes —, so hat man, da der Flügel in der Sekunde um die Strecke V vorrückt, einen sekundlichen Zuwachs an Bewegungsgröße von $\varrho V \int l\, d\Gamma$ (das Integral genommen über die Wirbelfäden eines Vorzeichens), der gleich dem Auftrieb des Flügels zu setzen ist. Es muß also im Wirbelschweif überall das $\int l\, d\Gamma$ konstant und gleich dem über den Tragflügel genommenen Integral $\int \Gamma\, d x$ sein.

16. Für die Beurteilung der Zustände im Wirbelschweif weit ab vom Tragwerk darf die Eigenbewegung der Wirbellinien, die für die Berechnung des Geschwindigkeitsfeldes am Tragwerk selbst vernachlässigt wurde, nicht mehr unbeachtet bleiben. Diese Bewegung besteht einerseits in einem Aufrollen des oder der Wirbelbänder von den Flügelspitzen aus und — bei Mehrdeckern — in einem seilartigen Umeinanderschlingen der durch das Aufrollen erzeugten Wirbelstränge, andererseits in einer gemeinsamen fortschreitenden Bewegung des ganzen Wirbelsystems senkrecht zu seiner Längserstreckung. Die Wirbelstärke wird durch weitere Andauer der Aufrollbewegung in jedem kleinen Volumteil immer mehr vergleichmäßigt und strebt einer stetigen räumlichen Wirbelverteilung als Grenzwert zu. Die gesamte Wirbelstärke der einen Seite stimmt in normalen Fällen mit der Zirkulation in der Mitte des Tragwerks überein. Der in Nr. 6 der I. Mitteilung abgeleitete Satz, daß für die Bewegung relativ zu dem Tragwerk in dem Wirbelsystem die Bernoulli'sche Gleichung mit einer universellen Konstanten gilt und daß deshalb die Wirbellinien mit den Stromlinien relativ zum Tragwerk identisch sind, bleibt auch für diesen Grenzzustand in Kraft.

Im Fall des Eindeckers erhält man zwei gerade Wirbelfäden, bei denen allerdings, wenn nicht eben die Auftriebsdichte über den größten Teil des Flügels konstant ist, auch außerhalb der eigentlichen Wirbelkerne schwache Rotation vorhanden ist. Nimmt man zur Vereinfachung an, daß alle Rotation in zwei Wirbelkernen von kleinem Durchmesser vereinigt ist, so erhält man außerhalb der Wirbelkerne eine ebene Bewegung. Innerhalb der Wirbelkerne ist wegen der schraubenförmigen Gestalt der Wirbellinien

auch eine axiale Komponente v' in der Richtung von dem Flügel weg vorhanden[1]).

Für die nachfolgende Betrachtung über den Widerstand interessiert der mit dieser Bewegung verknüpfte Unterdruck. Für einen Punkt außerhalb der Wirbelkerne, der von den Wirbelachsen die Entfernungen r_1 und r_2 aufweist, ist dieser, wie eine leichte Rechnung lehrt, durch die Gleichung gegeben:

$$p_0 - p = \frac{\varrho(u^2+w^2)}{2} = \frac{\varrho \Gamma_0^2}{8\pi^2} \cdot \frac{l^2}{r_1^2 r_2^2} \qquad (61)$$

wobei l den Abstand der beiden Wirbelfäden und Γ_0 deren Zirkulation bedeutet. Für das Innere der Wirbelkerne wird, wenn der Einfachheit halber die übliche Annahme gemacht wird, daß die Rotation im Innern jedes Wirbelkernes konstant ist, der Unterdruck $= \varrho u'^2 \left(1 - \frac{1}{2}\left(\frac{r}{r'}\right)^2\right)$, wobei r' der Kernradius und $u' = \Gamma_0/2\pi r'$ die Umfangsgeschwindigkeit im Radius r' ist. Der Steigungswinkel der schraubenförmigen Wirbellinien auf der Mantelfläche der Kurve ist gegeben durch $\operatorname{tg}\alpha = v/u'$, die lineare Steigung der Schraubenlinien ist daher

$$H = 2\pi r' V/u' = 4\pi^2 r'^2 V/\Gamma_0. \qquad (62)$$

Die Axialgeschwindigkeit v' ergibt sich aus der Schraubengestalt der Wirbellinien oder auch aus der Bernoulli'schen Gleichung genähert zu

$$v' = \frac{u'^2}{V}\left(1 - \left(\frac{r}{r'}\right)^2\right). \qquad (63)$$

Das ganze Wirbelsystem bewegt sich mit der Geschwindigkeit

$$w' = \Gamma_0/2\pi l \qquad (64)$$

abwärts.

Bei einem Doppeldecker ergeben sich vier Wirbelfäden, die zu zweien je als Rechtsschraube (rechts) und Linksschraube (links) verschlungen sind. Wird angenommen, daß die Entfernung der zwei Wirbel einer Schraube gleich der Doppeldeckerhöhe h ist, und daß beide die gleiche Stärke Γ_1 haben, so ergibt sich die Geschwindigkeit, mit der die beiden Wirbel umeinander umlaufen, angenähert zu $u_1 = \Gamma_1/2\pi h$. Damit wird die Schraubensteigung

1) Die Komponente der Gesamtgeschwindigkeit ist demnach $V + v + v'$, wo v die der Gl. (31) entsprechende Komponente ist. Diese ist in der Gegend der y-Achse sehr klein und nimmt überdies mit der Entfernung vom Tragwerk beliebig ab, kann deshalb hier außer Acht bleiben.

$$H = 2\pi \frac{h}{2} \cdot V / u_1 = 2\pi^2 h^2 V / \Gamma_1. \qquad (62\,a)$$

Die Geschwindigkeit v' ist in der Mitte der Schraube genähert

$$v' = \frac{2\Gamma_1}{2\pi h/2} \cdot \frac{u_1}{V} = \frac{\Gamma_1^2}{\pi^2 h^2 V}. \qquad (63\,a)$$

17. Der Widerstand hängt nach Früherem eng mit der kinetischen Energie des Wirbelsystems zusammen. Diese bleibt unter der Voraussetzung der Reibungslosigkeit bei der Umformung des Wirbelsystems unverändert. Ihr Betrag für die Längeneinheit, der mit dem des Widerstandes übereinstimmt, ist — unter Vernachlässigung des von höherer Ordnung kleinen v' —

$$W = \frac{\varrho}{2} \int\!\!\int (u^2 + w^2)\, dF. \qquad (65)$$

In einem Schnitt, der senkrecht zur Achse des Wirbelsystems geführt wird, finden sich die in voriger Nummer beschriebenen Unterdrücke, die als Resultante eine Saugkraft ergeben. Diese stimmt nun mit dem obigen Wert für den Widerstand nicht überein, sondern ist größer als dieser[1]). Diese Differenz wird durch den Impuls der Geschwindigkeit v' ausgeglichen, der der Saugkraft entgegenwirkt. Um dies klar zu stellen, wenden wir den Impulssatz für die stationäre Bewegung auf die Komponenten parallel der Achse des Wirbelsystems an und benützen hierfür eine Kontrollfläche, die aus einer Ebene senkrecht zu dem Wirbelsystem weitab vom Flügel und einer zu ihr parallelen Fläche weit vor dem Flügel besteht, mit einem im übrigen beliebigen Abschluß nach den Seiten und nach oben und unten, von dem nur angenommen wird, daß er sehr weit vom Flügel abliegt. Dann ist für ein Flächenelement dF der hinteren Kontrollebene die Saugkraft $= (p_0 - p)\, dF$ und der longitudinale Impuls $= dm(V - (V + v')) = -\varrho\,(V + v')\, dF \,.\, v'$. Die Bernoulli'sche Gleichung liefert:

$$p_0 - p = \varrho V v' + \frac{\varrho}{2} (v'^2 + u^2 + w^2).$$

Die Gesamtwirkung von Saugkraft und Impuls auf dem Flächenelement dF ergibt demnach

$$dW = dF \cdot \frac{\varrho}{2} (u^2 + w^2 - v'^2)\,^{2)}. \qquad (66)$$

1) Beim Eindecker z. B. beträgt die Differenz $= \frac{\varrho \Gamma_0^2}{4\pi} =$ rd $^1/_5$ des Widerstandes.

2) Zu derselben Formel wird man geführt, wenn man die Energie des se-

Streng genommen muß noch berücksichtigt werden, daß wegen der Abwärtsbewegung des Wirbelsystems mit der Geschwindigkeit w' die Achse des Wirbelsystems gegen die Flugrichtung um einen Winkel δ nach unten geneigt ist, für den gilt: $\sin\delta = w'/V$ (wobei w' senkrecht zur Wirbelachse gemessen ist). Man findet bei einer genaueren Betrachtung, daß die Resultante der aus der obigen Überlegung folgenden Kraft in die Richtung der Achse des Wirbelsystems fällt, also ebenfalls unter dem Winkel δ geneigt ist. Der Widerstand ist dann die Projektion dieser Kraft auf die Y-Achse, so daß also die Beziehung zwischen dem Widerstand und den Zustandsgrößen im Wirbelschweif in strenger Form lautet:

$$W = \frac{\varrho}{2} \cos\delta \iint (u^2 + w^2 - v'^2)\, dF\,[1]. \qquad (67)$$

Andererseits ergibt sich aus derselben Betrachtung der endgültige Wert des Auftriebes zu

$$A = \varrho V \int \Gamma\, dx - W \operatorname{tg} \delta. \qquad (68)$$

Das Korrektionsglied $W \operatorname{tg} \delta$ ist übrigens, wenn A von der I. Ordnung klein ist, von der III. Ordnung, kann also praktisch immer außer acht bleiben.

18. Die Gleichung (65) erlaubt in einfachen Fällen, die Größe der Wirbelkerne abzuschätzen, zu denen sich das Wirbelband des Flügels zusammenrollt. Für den Eindecker kann unter der Annahme, daß die Wirbel sämtlich in zwei Kernen von gleichförmig verteilter Rotation enthalten sind, das Integral von Gl. (64) entsprechend den Angaben von Nr. 16 ausgewertet werden. Es liefert

$$W = \frac{\varrho \Gamma_0^2}{2\pi}\left(\ln \frac{l}{r'} + \frac{1}{4}\right). \qquad (69)$$

Nimmt man am Tragflügel elliptische Auftriebsverteilung an, so wird andererseits $W = 2A^2/\pi \varrho V^2 b^2$; die Zirkulation eines Wirbelstranges ergibt sich gleich der Zirkulation in der Mitte des Tragflügels $\Gamma_0 = \frac{4}{\pi} \frac{A}{\varrho V b}$; hiermit wird

---

kundlich neu gebildeten Stückes des Wirbelschweifes betrachtet. Dieser besteht aus der kinetischen Energie: $dF(V + v') \cdot \frac{\varrho}{2}(u^2 + w^2 + v'^2)$ und aus der Arbeitsleistung des Druckes $dF(p - p_0)v'$, was unter Benutzung der Bernoulli'schen Gleichung wieder den obigen Ausdruck gibt.

1) Dies ist auch energetisch einzusehen, da das sekundlich neugebildete Stück des Wirbelschweifes nicht die Länge V, sondern wegen seiner Eigenbewegung nur $V \cos\delta$ hat.

$$\ln \frac{l}{r'} = \frac{1}{4}(\pi^2 - 1) = 2{,}215, \text{ also } l/r' = 9{,}16.$$

Die am Schluß von Nr. 15 angeführte Beziehung liefert hier $l = \frac{\pi}{4} b$, so daß unter der Voraussetzung der obigen Annahmen sämtliche Abmessungen des Wirbelsystems bekannt sind. Die Steigung der Stromlinien im Kern wird gemäß Gl. (62), wenn c_a und b anstelle von Γ_0, l und r' eingeführt werden:

$$H = 0{,}455\, b^3 / c_a F. \tag{70}$$

Die Geschwindigkeit in der Ebene des Wirbelpaares wird außerhalb der Wirbelkerne

$$w = 2\Gamma_0 l / \pi (l^2 - 4x^2), \tag{71}$$

innerhalb der Kerne nach unserer Annahme eine lineare Funktion. Im Gegensatz zu der Geschwindigkeit ohne Aufrollen des Wirbelbandes $w_1 = 2w_0 = 4A/\pi \varrho V b^2$ wird demnach die Geschwindigkeit in der Mitte des Wirbelpaares

$$w_2 = 4w' = 2\Gamma_0 / \pi l = 8\, w_1 / \pi^2. \tag{72}$$

Fig. 4 gibt ein Bild der beiden Geschwindigkeitsverteilungen auf

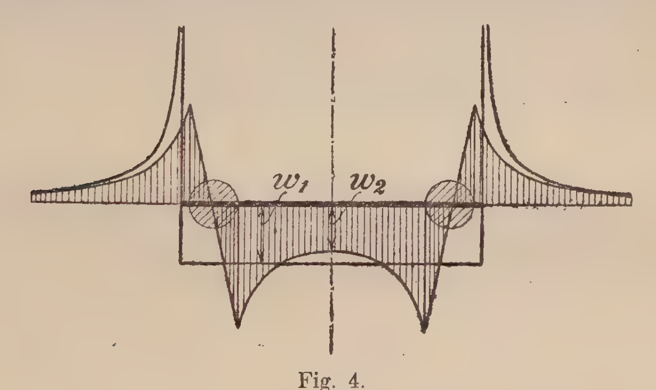


Fig. 4.

den Parallelen zur X-Achse, wobei noch erwähnt werden mag, daß für das nicht aufgerollte Wirbelband außerhalb des Tragflügels $w_1' = -w_1 \left(\frac{x}{\sqrt{x^2 - (l/2)^2}} - 1 \right)$ ist.

Die Entfernung vom Flügel, in welcher sich das Aufrollen der Hauptsache nach vollzieht, ist umso größer, je kleiner die Geschwindigkeit w_1 ist; denn sie ist offenbar von der Größenordnung $\frac{1}{2} l V / w_1$. Die genauere Aufklärung des Aufrollvorgangs hat —

wegen der Verschiedenheit von w_1 und w_2 — eine gewisse Bedeutung für die Berechnung der Zustände an einem verhältnismäßig weit hinter dem Tragflügel befindlichen Höhenleitwerk (Steuerfläche).

Die in den beiden Mitteilungen dargelegte Flügeltheorie ist nach verschiedenen Richtungen hin der Ergänzung und Erweiterung fähig. Einerseits läßt sich durch Übergang von der tragenden Linie zur tragenden Fläche eine Verfeinerung gewinnen, die nicht nur die Einflüsse der Profilgestalt auf den Auftrieb und das Moment zu untersuchen gestattet, sondern auch den bei Doppeldeckern usw. sehr bemerkbaren Wölbungseinfluß besser als bisher zu berechnen erlaubt; ferner lassen sich auch die Verhältnisse bei schräggestellten und in der Flugrichtung gekrümmten Flügeln mit solchen Ansätzen klären. Eine genäherte Ermittlung der Wirkungen der zeitlichen Änderung des Auftriebs, z. B. beim Flügelschlag, ist mit ihnen ebenfalls möglich. Andererseits lassen sich die Ansätze der Flügeltheorie auch auf den Kurvenflug, sowie auf die Vorgänge an Luftschrauben übertragen, wobei im Hintergrunde noch die Anwendung auf Gebläseräder und Turbinen steht. Neben den obengenannten Problemen harren auch auf den in den beiden Mitteilungen behandelten Gebieten noch zahlreiche Aufgaben der Durchführung.

Mit der Bearbeitung verschiedener dieser Aufgaben ist bereits begonnen worden. Über ihre Ergebnisse hoffe ich später an dieser Stelle zu berichten. Im übrigen liegt es mir fern, mit den obigen Bemerkungen das ganze Arbeitsgebiet für Göttingen reservieren zu wollen; im Gegenteil werde ich die Mitarbeit der Fachgenossen jederzeit gerne begrüßen. Nur würde in diesem Fall zur Vermeidung von Doppelarbeit eine vorherige gegenseitige Bezugnahme zweckmäßig sein.

Zusammenfassung. I. Mitteilung: Die Hydrodynamik der reibungslosen Flüssigkeit wird durch Hinzufügung von Sätzen über die Entstehung von Wirbeln am festen Körper ergänzt, und eine Tragflügeltheorie entwickelt, bei der die Flügel durch geeignete Wirbelgebilde ersetzt werden. Es wird eine „Theorie der I. Ordnung" aufgestellt, in der unter Annahme kleiner Störungsgeschwindigkeiten bei jeder Zustandsgröße nur die jeweils niedrigste Ordnung berücksichtigt wird. Durch diese Theorie werden die Zusammenhänge zwischen der Verteilung des Auftriebes über die

Spannweite und dem Anstellwinkel und dem Widerstande aufgeklärt; im besondern werden die auf den geraden Eindecker bezüglichen Formeln aufgestellt.

II. Mitteilung: Die Theorie der I. Ordnung wird auf beliebige Flügelsysteme („Tragwerke“) erweitert, und u. a. ein neuer Beweis des Munk'schen Satzes gebracht, daß auf den Gesamtwiderstand eines Tragwerks eine Änderung der „Staffelung“ ohne Einfluß bleibt, wenn die Auftriebsverteilung der einzelnen Teile dabei ungeändert bleibt. Die Ergebnisse der Munk'schen Dissertation über die Tragwerke kleinsten Widerstandes werden auf anderem Wege hergeleitet und durch eigene Betrachtungen ergänzt. Die bei den Kanälen und freien Luftstrahlen der Versuchsanstalten durch die Begrenzung des Luftstromes auftretenden Fehlereinflüsse werden dargelegt und berechnet. Die Zustände in großer Entfernung vom Tragwerk und im Wirbelschweif werden erörtert und ihr Zusammenhang mit dem Auftrieb und dem Widerstand klargelegt. — Ein Ausblick auf die weiteren Aufgaben der Theorie wird angefügt.

Die Endlichkeit des Systems der ganzzahligen Invarianten binärer Formen.

Von

Emmy Noether in Göttingen.

Vorgelegt von F. Klein in der Sitzung vom 27. März 1919.

Im folgenden wird, als Erweiterung des bekannten Endlichkeitssatzes, der Satz bewiesen: Alle ganzen **ganzzahligen** Invarianten eines binären Grundformensystems sind ganze rationale, **ganzzahlige** Funktionen einer endlichen Anzahl unter ihnen. Dabei ist wie üblich, unter „Invariante" immer eine ganze rationale Invariante einer Grundform oder eines Grundformensystems zu verstehen, und zwar eine Invariante gegenüber der durch die Gruppe aller linearen Transformationen induzierten Koeffiziententransformation.

Der Beweis beruht auf einer Ganzzahligkeits-Verschärfung des Mertens-Hilbertschen Endlichkeitsbeweises[1]). Dieser Endlichkeitsbeweis läßt sich so charakterisieren: Man zerlegt die Grundformen in ihre Linearfaktoren — präziser ausgedrückt: man ordnet jeder Grundform ein Produkt von Linearformen zu —, wodurch jede Invariante I in eine Invariante dieser Linearformen, also in eine Funktion der homogen gemachten Wurzeln übergeht. Diese Funktion hat drei Bedingungen zu genügen: 1) Der Bedingung der Invarianz, 2) Gewichtsbedingungen, 3) Symmetriebedingungen. Die Bedingungen 2) und 3) sind notwendig, damit I ganz und rational in den Koeffizienten der Grundformen wird. Der Bedin-

---

1) F. Mertens, Wiener Sitzungsberichte Jan. 1889; D. Hilbert, Math. Ann. Bd. 33 (1889) [datiert März 1888]. Die beiden, unabhängig von einander im Anschluß an Mertens, Crelle Bd. 100 entstandenen Beweise stimmen inhaltlich überein.

gung 1) wird dadurch genügt, daß I als ganze rationale Funktion der Determinanten der Linearformen dargestellt wird; der Bedingung 2) dadurch, daß statt der einzelnen Determinanten gewisse ihrer Potenzprodukte als neue Argumente genommmen werden. Die Bedingung 3) sagt jetzt einfach aus, daß I eine ganze rationale, symmetrische Funktion von Größenreihen wird, und daß umgekehrt jede solche symmetrische Fumktion von Größenreihen eine Invariante der Grundformen wird. Die symmetrischen Elementarfunktionen dieser Größenreihen ergeben also das vollständige System.

Betrachtet man nun nur ganzzahlige Invarianten, so läßt sich der Bedingung 1) ebenso wie oben genügen, sobald für Invarianten von Linearformen der schärfere Satz bewiesen ist, daß alle **ganzzahligen** Invarianten binärer Linearformen ganze **ganzzahlige** Funktionen der Determinanten dieser Linearformen sind. Daß dies tatsächlich zutrifft, zeige ich in § 3, und zwar in der Fassung, daß die Gesamtheit aller modulo irgend einer Primzahl zwischen diesen Determinanten bestehenden Relationen übereinstimmt mit der Gesamtheit aller identischen Relationen; daß also der diesen Relationen entsprechende Modul ein Primmodul modulo jeder Primzahl bleibt. Dieser Nachweis geschieht durch Normierung der Restklassen inbezug auf den Modul der quadratischen Formen, die den quadratischen Relationen zwischen den Determinanten entsprechen; diese quadratischen Formen ergeben sich dabei als ganzzahlige Basis des Moduls, und zwar modulo jeder Primzahl; während sie bisher nur unter Vernachlässigung der Ganzzahligkeit als Modulbasis bekannt waren.

Um der Bedingung 2) zu genügen, ist keine Abänderung nötig, da die Einführung der Potenzprodukte die Zahlkoeffizienten nicht beeinflußt. Die Bedingung 3) sagt jetzt aus, daß $n_1! \dots n_\mu! I$ (wo $n_1, \dots n_\mu$ die Gradzahlen der einzelnen Grundformen bedeuten) eine ganzzahlige symmetrische Funktion von Größenreihen wird, und daß umgekehrt jede ganzzahlige symmetrische Funktion dieser Reihen eine ganzzahlige Invariante wird. Nun besitzen aber, wie sich leicht zeigen läßt, diese ganzzahligen symmetrischen Funktionen von Reihen eine ganzzahlige Basis, zu der allerdings die Elementarfunktionen nicht mehr ausreichen. Die Anwendung des Hilbertschen Theorems von der ganzzahligen Modulbasis ergibt daraus auch für die I selbst eine ganzzahlige Basis.

§ 1 stellt die den Bedingungen 1), 2), 3) entsprechenden speziellen Endlichkeitssätze zusammen, woraus dann in § 2 der End-

lichkeitsbeweis folgt. Die der Bedingung 3) entsprechenden speziellen Endlichkeitssätze ergeben zugleich die Endlichkeit der ganzzahligen Invarianten endlicher Gruppen ganzzahliger oder algebraisch-ganzzahliger Substitutionen, wie in § 4 noch kurz ausgeführt wird.

§ 1. Spezielle Endlichkeitssätze.

Ich schicke die folgenden speziellen Endlichkeitssätze voraus, die dazu dienen werden, den in der Einleitung erwähnten Bedingungen der Invarianz, des Gewichts und der Symmetrie zu genügen:

I. Jede ganzzahlige Invariante eines Systems binärer Linearformen ist eine ganze rationale, ganzzahlige Funktion der Determinanten dieser Linearformen (der Beweis wird in § 3 erbracht werden).

II. Jedes System $\mathfrak{S}$ von Potenzprodukten von n Variabeln mit nichtnegativen Exponenten, das neben irgend zwei Potenzprodukten f und g auch deren Quotienten enthält, sobald dieser Quotient ganz in den Variabeln ist, besitzt eine endliche multiplikative Basis $f_1, \dots f_k$ von Funktionen aus $\mathfrak{S}$, derart daß für jedes f aus $\mathfrak{S}$ eine Darstellung gilt: $f = f_1^{\lambda_1} \dots f_k^{\lambda_k}$, mit nicht-negativen Exponenten $\lambda_1 \dots \lambda_k$.

Nach dem Hilbertschen Satz von der Modulbasis existiert nämlich eine endliche Anzahl von Funktionen aus $\mathfrak{S}$: $f_1, \dots f_k$, die sämtlich die x wirklich enthalten, derart daß jedes nicht-konstante f aus $\mathfrak{S}$ dem daraus abgeleiteten Modul angehört: $f \equiv 0 \mod (f_1 \dots f_k)$. Daraus folgt, daß zu jedem $f = x_1^{i_1} \dots x_n^{i_n}$ mindestens ein $f_\nu = x_1^{\nu_1} \dots x_n^{\nu_n}$ existiert derart daß $\nu_1 \leqq i_1; \dots \nu_n \leqq i_n$. In der Darstellung:

$$f = x_1^{i_1 - \nu_1} \dots x_n^{i_n - \nu_n} f_\nu = A f_\nu \text{ [1]}$$

1) Diese Darstellung läßt sich auch direkt beweisen, und daraus folgt dann umgekehrt wieder der Hilbertsche Satz: vergl. P. Gordan: Neuer Beweis des Hilbertschen Satzes über homogene Funktionen, Gött. Nachr. 1899. Ein anderer direkter Beweis dieser Darstellung, aus der dann auch Satz II abgeleitet wird, bei A. Ostrowski: Über die Existenz einer endlichen Basis bei gewissen Funktionensystemen (Math. Ann. 78, 1916) § 2, 1 und § 3, 2.

Der in § 2 benutzte Spezialfall von Satz II, wo die Exponenten alle nichtnegativen Lösungen eines Diophantischen Gleichungssystems durchlaufen, findet sich schon bei Gordan-Kerschensteiner: Vorlesungen über Invariantentheorie, Bd. I, S. 199.

gehört nach Voraussetzung A zu $\mathfrak{S}$, hat aber niedrigeren Grad in den Variabeln als f, so daß durch endlich oftmalige Wiederholung der Schlußweise die Behauptung folgt (damit die Darstellung auch noch für $f = 1$ gilt, sind nur alle Exponenten λ gleich Null zu setzen).

III. Alle ganzen ganzzahligen symmetrischen Funktionen von n Größenreihen sind ganze, ganzzahlige Funktionen einer endlichen Anzahl unter ihnen.

Denn bezeichnen $\tau_1(x), \ldots \tau_n(x)$ die elementaren symmetrischen Funktionen der Unbestimmten $x_1 \ldots x_n$, so gilt für jeden Exponenten k eine Identität:

$$x_i^k = G_0^{(k)}(\tau) + x_i G_1^{(k)}(\tau) + \cdots + x_i^{n-1} G_{n-1}^{(k)}(\tau), \quad (i = 1, 2 \ldots n),$$

wo $G_0^{(k)}, \ldots G_{n-1}^{(k)}$ ganze, ganzzahlige Funktionen der $\tau(x)$ bedeuten. Es seien nun $x_i, y_i \ldots z_i$ die Elemente der $i$ten Reihe; dann werden vermöge dieser Identitäten und entsprechender für die $y \ldots z$ alle speziellen (einförmigen) symmetrischen Funktionen:

$$\sum_{i=1}^{n} x_i^\alpha y_i^\beta \ldots z_i^\gamma$$

ganze ganzzahlige Funktionen der endlich vielen:

$$\sum_{i=1}^{n} x_i^a y_i^b \ldots z_i^c \quad (0 \leqq a < n;\ 0 \leqq b < n;\ \ldots 0 \leqq c < n)$$

und der elementaren symmetrischen Funktionen $\tau(x), \tau(y) \ldots \tau(z)$[1].

Hat man den im folgenden nicht auftretenden Fall der allgemeinsten symmetrischen Funktionen der n Reihen, so zeigt sich ebenso, daß die Funktionen

$$\sum x_{i_1}^{a_1} y_{i_1}^{b_1} \ldots z_{i_1}^{c_1} x_{i_2}^{a_2} y_{i_2}^{b_2} \ldots z_{i_2}^{c_2} \ldots x_{i_n}^{a_n} y_{i_n}^{b_n} \ldots z_{i_n}^{c_n}$$
$$(0 \leqq a_\lambda < n;\ 0 \leqq b_\lambda < n;\ \ldots 0 \leqq c_\lambda < n),$$

wo $i_1, i_2 \ldots i_n$ alle Permutationen durchläuft, zusammen mit den $\tau(x), \tau(y) \ldots \tau(z)$ eine ganzzahlige Basis bilden.

IV. Es seien $F_1(x), \ldots F_k(x)$ ganze ganzzahlige Funktionen von $x_1 \ldots x_n$; a eine feste ganze Zahl. Dann sind alle Funktionen $\frac{G(F)}{a}$, wo G eine ganze ganzzahlige Funktion bedeutet, die ganzzahlig in den

1) Hilbert nimmt a. a. O. anstelle der $\tau(x), \ldots \tau(z)$ die Potenzsummen in die Basis auf, wodurch auch diese nur aus einförmigen Funktionen besteht, aber die Ganzzahligkeit der Darstellung verloren geht.

x werden, ganze ganzzahlige Funktionen einer endlichen Anzahl unter ihnen.

Denn nach dem Hilbertschen Satz von der ganzzahligen Modulbasis gilt für jedes solche $\frac{G(F)}{a}$ eine Darstellung:

$$\frac{G(F)}{a} = A_1(F_1 \cdots F_k)\frac{G_1(F)}{a} + \cdots + A_\varrho(F_1 \cdots F_k)\frac{G_\varrho(F)}{a},$$

wo $A_1, \ldots A_\varrho$ ganze ganzzahlige Funktionen der F bedeuten, und $\frac{G_1(F)}{a}, \cdots \frac{G_\varrho(F)}{a}$ zum System gehören. Die Funktionen

$$F_1 = \frac{aF_1}{a}; \cdots F_k = \frac{aF_k}{a}; \qquad \frac{G_1(F)}{a}, \cdots \frac{G_\varrho(F)}{a}$$

bilden also die gewünschte Basis.

§ 2. Endlichkeit der ganzzahligen binären Invarianten.

Es seien x, y binäre Veränderliche, die einer Transformation mit unbestimmten Koeffizienten

(1) $$x = s_{11}x' + s_{12}y'; \qquad y = s_{21}x' + s_{22}y'$$

unterworfen werden; f eine binäre Grundform, deren Koeffizienten Unbestimmte seien. Dann gilt vermöge (1):

(2) $$f = a_0x^n + a_1x^{n-1}y + \cdots + a_ny^n = a_0'x_n'^n + a_1'x'^{n-1}y' + \cdots + a_n'y'^n.$$

Unter einer ganzzahligen Invariante von f versteht man bekanntlich eine ganze ganzzahlige Funktion $I(a)$, derart daß

(3) $$I(a') = |s_{ik}|^\varrho I(a), \qquad \text{(identisch in } a, s_{ik})^{1)};$$

entsprechend sind die ganzzahligen Invarianten eines Grundformensystems zu definieren, wobei noch die Homogenität in den Koeffizienten der einzelnen Grundformen mit in die Definition aufgenommen werden kann, da aus solchen einzeln-homogenen Invarianten sich die übrigen ganzzahlig-linear zusammensetzen lassen.

Ich betrachte nun neben (2) noch eine spezielle binäre Form g, nämlich das Produkt von n Linearformen, deren Koeffizienten wieder Unbestimmte seien:

---

1) Schreibt man, wie vielfach in der Invariantentheorie üblich, f mit Binomialkoeffizienten: $f = b_0x^n + \binom{n}{1}b_1x^{n-1}y + \cdots + b_ny^n$, so wird der Bereich aller ganzzahligen Invarianten $I(b)$ gegenüber dem der $I(a)$ erweitert, und die angewandten Endlichkeitsschlüsse versagen; $I(b)$ wird keine ganzzahlige Funktion der Wurzeln mehr.

$$(4) \qquad \begin{aligned} g &= (\alpha_1 x + \beta_1 y)(\alpha_2 x + \beta_2 y) \ldots (\alpha_n x + \beta_n y) \\ &= \sigma_0(\alpha, \beta) x^n + \sigma_1(\alpha, \beta) x^{n-1} y + \cdots \sigma_n(\alpha, \beta) y^n \\ &= (\alpha_1' x' + \beta_1' y')(\alpha_2' x' + \beta_2' y') \ldots (\alpha_n' x' + \beta_n' y') \\ &= \sigma_0' x'^n + \sigma_1' x'^{n-1} y' + \cdots \sigma_n' y'^n, \end{aligned}$$

wo also $\sigma_0(\alpha, \beta)$, $\sigma_1(\alpha, \beta) \ldots \sigma_n(\alpha, \beta)$ die homogen gemachten symmetrischen Elementarfunktionen bedeuten, und die transformierten Koeffizienten σ_i' gleich $\sigma_i(\alpha', \beta')$ werden.

Wegen der algebraischen Unabhängigkeit dieser Elementarfunktionen entspricht jeder Relation:

$$(5) \qquad G(\sigma_0, \sigma_1 \ldots \sigma_n) = 0 \quad [\text{identisch in } \alpha, \beta]$$

eine weitere:

$$(6) \qquad G(\sigma_0, \sigma_1 \ldots \sigma_n) = 0 \quad [\text{identisch in } \sigma_0, \ldots \sigma_n],$$

und somit für Unbestimmte $a_0, \ldots a_n$:

$$(7) \qquad G(a_0, a_1 \ldots a_n) = 0 \quad [\text{identisch in } a_0, \ldots a_n].$$

Nun entspricht jeder ganzzahligen Invariante $I(a)$ von f durch die Substitution $a_i = \sigma_i(\alpha, \beta)$ eine ganzzahlige Invariante $I(\sigma) = H(\alpha, \beta)$ von g, wo H eine ganzzahlige Funktion von α, β wird, für diese Invariante geht die Relation (3) über in:

$$(8) \qquad I(\sigma') = |s_{ik}|^{\varrho} I(\sigma) \quad [\text{identisch in } \sigma, s_{ik}],$$

und wegen $\sigma' = \sigma(\alpha', \beta')$ folgt aus (8):

$$(9) \qquad H(\alpha', \beta') = |s_{ik}|^{\varrho} H(\alpha, \beta) \quad [\text{identisch in } \alpha, \beta;\ s_{ik}];$$

$I(\sigma) = H(\alpha, \beta)$ wird also eine ganzzahlige Invariante der n Linearformen $(\alpha_i x + \beta_i y)$. Es sei nun umgekehrt $H(\alpha, \beta)$ eine ganzzahlige Invariante dieser n Linearformen, also (9) erfüllt, und es sei zugleich $H(\alpha, \beta)$ gleich einer ganzzahligen Funktion $I(\sigma)$ der σ. Da dann auch

$$H(\alpha', \beta) = I(\sigma(\alpha', \beta')) = I(\sigma')$$

wird, schreibt sich (9) auch:

$$(9\text{a}) \qquad I(\sigma') = |s_{ik}|^{\varrho} I(\sigma) \quad [\text{identisch in } \alpha, \beta;\ s_{ik}];$$

und daraus folgt nach (5) und (6) auch (8); und nach (7) auch (3). Es entspricht somit jeder ganzzahligen Invariante $I(a)$ eine ganzzahlige Invariante $H(\alpha, \beta) = I(\sigma)$ der n Linearformen, und umgekehrt. Bedeutet ferner $I_1(a), \ldots I_k(a)$ eine ganzzahlige Basis aller $I(\alpha)$, so wird $H_1(\alpha, \beta), \ldots H_k(\alpha, \beta)$ eine ganzzahlige Basis aller $H(\alpha, \beta)$, die zugleich ganzzahlige

Funktionen $I(\sigma)$ sind; und auch hier gilt nach (5), (6), (7) die Umkehrung. Um also die Existenz einer ganzzahligen Basis für alle $I(a)$ nachzuweisen, genügt der Nachweis für alle ganzzahligen Invarianten $H(\alpha, \beta)$, die zugleich ganzzahlige Funktionen $I(\sigma)$ sind; und die Basis $H_1(\alpha, \beta) = I_1(\sigma), \ldots H_k(\alpha, \beta) = I_k(\sigma)$ liefert für die $I(a)$ die Basis $I_1(a), \ldots I_k(a)$.

Es durchlaufe nun $H(\alpha, \beta) = I(\sigma)$ die Gesamtheit aller ganzzahligen Invarianten von (4). Da wegen (3) jede Invariante homogen in den σ wird, und nach (4) jedes σ homogen und linear in den Koeffizienten jeder Linearform ist, wird also $H(\alpha, \beta)$ homogen in jeder Reihe $\alpha_i \beta_i$ und in jeder Reihe von gleicher Dimension; außerdem natürlich eine symmetrische Funktion der n Reihen. Umgekehrt wird jede ganzzahlige symmetrische Funktion $H(\alpha, \beta)$ der n Reihen, die einzeln homogen und gleicher Dimension in jeder Reihe ist, nach dem Fundamentalsatz über die symmetrischen Funktionen eine ganze, ganzzahlige Funktion der σ; also nach dem obigen eine Invariante $I(\sigma)$, sobald $H(\alpha, \beta)$ invariant ist. Die Gesamtheit der ganzzahligen Invarianten von (4) ist also charakterisiert als Gesamtheit der ganzzahligen Funktionen $H(\alpha, \beta)$ mit folgenden Eigenschaften:

1) Ganzzahlige Invariante der n Linearformen $\alpha_i x + \beta_i y$;
2) Einzeln homogen und gleicher Dimension in jeder der n Reihen α_i, β_i;
3) Symmetrisch in den n Reihen α_i, β_i.

Nach dem speziellen Endlichkeits-Satz I wird $H(\alpha, \beta)$ wegen 1) eine ganze, ganzzahlige Funktion der Determinanten

$$(ik) = \alpha_i \beta_k - \beta_i \alpha_k;$$

also

$$H(\alpha, \beta) = G((ik)).$$

In dieser Darstellung läßt sich die Symmetrie auch formal zum Ausdruck bringen, indem man auf $G((ik))$ alle Permutationen der n Reihen anwendet:

$$(10) \quad n!\, H(\alpha, \beta) = G^{(1)}((ik)) + G^{(2)}((ik)) + \cdots + G^{(n!)}((ik)) = G^*((ik)).$$

G^* enthält also neben jedem Potenzprodukt P der (ik) auch die Summe $Q = P^{(1)} + \cdots + P^{(n!)}$ über alle Permutationen von P; und zwar wird G^* eine ganzzahlige, lineare Kombination endlich vieler Q, $G^* = \sum c Q$. Da die Q den Bedingungen 1), 2), 3) genügen, also selbst Invarianten werden, genügt es die Q anstelle von $G^*((ik))$ zu betrachten.

Es durchlaufe nun P das System $\mathfrak{S}$ aller in den Q auftretenden, d. h. der Bedingung 2) genügenden Potenzprodukte der (ik). Ist der Quotient zweier P ganz in den (ik), so genügt er ebenfalls der Bedingung 2), gehört also zu $\mathfrak{S}$. Nach Satz II gibt es somit eine endliche Anzahl derartiger P, etwa $P_1, \dots P_\varrho$ derart, daß jedes P sich darstellen läßt in der Form

$$P = P_1^{\lambda_1} \dots P_\varrho^{\lambda_\varrho}$$

mit nicht-negativen Exponenten λ. Durch Anwendung aller Permutationen erhält man daraus:

$$Q = P_1^{(1)\lambda_1} \dots P_\varrho^{(1)\lambda_\varrho} + \dots + P_1^{(n!)\lambda_1} \dots P_\varrho^{(n!)\lambda_\varrho};$$

Q wird also eine ganzzahlige symmetrische Funktion der $n!$ Größenreihen von je ϱ Elementen:

$$P_1^{(1)} \dots P_\varrho^{(1)}; \quad P_1^{(2)} \dots P_\varrho^{(2)}; \; \dots; \; P_1^{(n!)} \dots P_\varrho^{(n!)};$$

also nach Satz III eine ganze ganzzahlige Funktion endlich vieler symmetrischer Funktionen dieser Reihen. Diese Basisfunktionen werden selbst Größen Q oder symmetrische Elementarfunktionen der $n!$ Elemente $P_i^{(1)}, P_i^{(2)} \dots P_i^{(n!)}$, genügen somit den Bedingungen 1) 2) 3) und werden ganzzahlige Invarianten

(11) $$H_1(\alpha, \beta) = I_1(\sigma); \dots H_l(\alpha, \beta) = I_l(\sigma).$$

Nach (10) wird somit wegen $G^* = \sum c\,Q$ jedes $H(\alpha, \beta) = I(\sigma)$ von der Form:

$$H(\alpha, \beta) = I(\sigma) = \frac{1}{n!} \Phi(I_1, \dots I_l),$$

wo Φ eine ganze ganzzahlige Funktion bedeutet; und umgekehrt wird $\frac{1}{n!}\Phi$ eine ganzzahlige Invariante, sobald es ganz in α, β wird. Es lassen sich also nach Satz IV den Invarianten (11) noch endlich viele $I_{l+1}(\sigma), \dots I_k(\sigma)$ zufügen, derart daß

$$H_1(\alpha, \beta) = I_1(\sigma); \quad \dots H_l(\alpha, \beta) = I_l(\sigma);$$
$$H_{l+1}(\alpha, \beta) = I_{l+1}(\sigma); \dots H_k(\alpha, \beta) = I_k(\sigma)$$

eine ganzzahlige Basis aller Invarianten $H(\alpha, \beta) = I(\sigma)$ bilden, womit der Satz für den Fall einer Grundform bewiesen ist.

Handelt es sich um ein Grundformensystem, so kann, wie oben bemerkt, Homogenität der Invarianten in den Koeffizienten jeder Grundform vorausgesetzt werden. Man ersetzt nun wieder die allgemeinen Grundformen durch solche, die jeweils das Produkt

von Linearformen sind. Deren Invarianten werden dann: 1) Invarianten dieser Linearformen; 2) einzeln homogen in den aus den Koeffizienten der Linearformen gebildeten Reihen, und gleicher Dimension in den Reihen jeder einzelnen Grundform; 3) symmetrisch in den Reihen jeder einzelnen Grundform. Umgekehrt entspricht wieder jeder ganzzahligen Invariante aller Linearformen, die diesen Bedingungen genügt, eine ganzzahlige Invariante des Grundformensystems. Der Beweis verläuft dem angegebenen absolut parallel; auf jedes, den Gewichtsbedingungen genügende Potenzprodukt $P = P_1^{\lambda_1} \dots P_\varrho^{\lambda_\varrho}$ der einzelnen Determinanten (Satz I und II) sind alle $N = n_1!\, n_2! \dots n_\mu!$ Permutationen anzuwenden, wo $n_1, \dots n_\mu$ die Gradzahlen der Grundformen bedeuten. Dadurch geht $N.I$ über in eine ganzzahlige symmetrische Funktion von N Reihen aus je ϱ Elementen, woraus die ganzzahlige Basis für alle $N.I$ und folglich nach Satz IV die ganzzahlige Basis für alle I sich ergibt.

§ 3. Die Basis der ganzzahligen Invarianten binärer Linearformen.

Es ist jetzt noch der Beweis des speziellen Endlichkeitssatzes I nachzutragen. Nach einem bekannten, zuerst von Clebsch bewiesenen Satz sind alle ganzen rationalen Invarianten der n Linearformen $\alpha_i x + \beta_i y$ ganze rationale Funktionen der Determinanten (ik). Insbesondere werden also alle ganzzahligen Invarianten dieser Linearformen ganze rationale, aber nicht notwendig ganzzahlige Funktionen dieser Determinanten.

Im folgenden wird gezeigt, daß vermöge der zwischen den (ik) bestehenden identischen Relationen diese Darstellung auch stets ganzzahlig möglich ist.

In Formeln und in der Sprache der Modultheorie drückt sich das so aus: Es seien ξ_{ik}, wo $i < k$, Unbestimmte; dann bildet die Gesamtheit aller ganzzahligen Polynome $F(\xi_{ik})$ von der Eigenschaft, daß $F((ik))$ identisch in α, β verschwindet, einen Modul $\mathfrak{M}$; da die Summe zweier F und ebenso das Produkt eines F mit einem beliebigen ganzzahligen Polynom der ξ_{ik} auch dieser Gesamtheit angehört. In Formeln: aus

$$F((ik)) = 0 \quad [\text{identisch in } \alpha, \beta]$$

folgt

$$F(\xi_{ik}) \equiv 0 \bmod \mathfrak{M} \quad [\text{identisch in } \xi_{ik}],$$

und umgekehrt.

Ich zeige nun, daß ganz entsprechend gilt: Aus

$$F((ik)) \equiv 0 \bmod p \quad [\text{identisch in } \alpha, \beta]$$

(1) folgt

$$F(\xi_{ik}) \equiv 0 \bmod (\mathfrak{M}, p) \quad [\text{identisch in } \xi_{ik}],$$

und umgekehrt, und zwar für *jede* Primzahl p. Dadurch ist $\mathfrak{M}$ *also auch identisch mit dem der Gesamtheit aller Relationen der* (ik) *modulo* p *entsprechenden Modul* $\mathfrak{M}_p$, der aus allen ganzzahligen Polynomen $F(\xi_{ik})$ besteht, für die $F((ik))$ modulo p identisch in α, β verschwindet.

Daß (1) *die ganzzahlige Darstellbarkeit aussagt*, zeigt sich so: (1) läßt sich auch in der Form schreiben:

Aus $F((ik)) = p \cdot H(\alpha, \beta)$ [identisch in α, β] folgt:

$$F(\xi_{ik}) - p \cdot G(\xi_{ik}) \equiv 0 \bmod \mathfrak{M} \quad [\text{identisch in } \xi_{ik}],$$

also nach der Definition von $\mathfrak{M}$:

$$F((ik)) = p \cdot G((ik)) \quad [\text{identisch in } \alpha, \beta] \text{ [1]}.$$

Somit ergibt sich: $p \cdot H(\alpha, \beta) = p \cdot G((ik))$ oder: $H(\alpha, \beta) = G((ik))$.

Liegt also eine Darstellung vor:

$$H(\alpha, \beta) = \frac{1}{p_1^{\lambda_1} \cdots p_\varrho^{\lambda_\varrho}} F((ik)),$$

wo $p_1, \dots p_\varrho$ Primzahlen bedeuten, so folgt daraus auch:

$$H(\alpha, \beta) = \frac{1}{p_1^{\lambda_1 - 1} \cdots p_\varrho^{\lambda_\varrho}} G((ik)),$$

woraus durch endlich oftmalige Wiederholung sich die Ganzzahligkeit der Darstellung ergibt.

Zum Nachweis von (1) bezeichne

$$f_{ik,rs} = \xi_{ik}\xi_{rs} - \xi_{ir}\xi_{ks} + \xi_{is}\xi_{kr}, \qquad i < k < r < s$$

die den quadratischen Relationen zwischen den (ik):

$$(ik)(rs) - (ir)(ks) + (is)(kr) = 0$$

entsprechenden Formen der Unbestimmten ξ_{ik}; $\mathfrak{N} = (f_{ik,rs})$ den aus ihnen abgeleiteten Modul aus ganzzahligen Polynomen. Aus

$$F(\xi_{ik}) \equiv 0 \bmod \mathfrak{N} \quad \text{folgt also: } F((ik)) = 0 \qquad [\text{identisch in } \alpha, \beta],$$

(2) Aus

$$F(\xi_{ik}) \equiv 0 \bmod (\mathfrak{N}, p) \quad \text{„} \quad \text{„} \quad F((ik)) \equiv 0 \bmod p \; [\text{identisch in } \alpha, \beta].$$

---

1) Dieser letztere Schluß beweist zugleich die Umkehrung von (1).

Ich zeige nun, vermöge einer Normierung der Restklassen inbezug auf $\mathfrak{N}$, daß auch die *Umkehrung* von (2) gilt, womit $\mathfrak{N} = \mathfrak{M} = \mathfrak{M}_p$ und damit (1) bewiesen sein wird. Dabei ergibt sich $\mathfrak{N} = \mathfrak{M}$ als Spezialfall von $\mathfrak{N} = \mathfrak{M}_p$ für $p = 0$; *der Fall $p = 0$ ist in den folgenden Überlegungen stets mit enthalten.* Anders ausgedrückt: *Die $f_{ik,\,rs}$ bilden eine ganzzahlige Modulbasis von $\mathfrak{M}_p$, insbesondere auch von $\mathfrak{M}_0 = \mathfrak{M}$.*

a) *Fall von zwei oder drei Reihen* $\alpha_i \beta_i$. Da noch keine vier verschiedenen Indices auftreten, wird $\mathfrak{N}$ der Nullmodul. Es ist also zu zeigen, daß aus

(3) $$F((ik)) \equiv 0 \bmod p \ \text{[identisch in } \alpha, \beta] \text{ folgt:}$$
$$F(\xi_{ik}) \equiv 0 \bmod p \ \text{[identisch in } \xi_{ik}],$$

womit (1) mit $\mathfrak{M} = \mathfrak{M}_p = 0$ bewiesen sein wird.

Da die Bedingung $F(ik) = p\,.\,H(\alpha, \beta)$ für die in den einzelnen Reihen *homogenen* Bestandteile einzeln erfüllt sein muß, genügt es durchweg, $F((ik))$ *in jeder einzelnen Reihe homogen*, also $F(\xi_{ik})$ *in jedem Index homogen*, anzunehmen. Nun gibt es im Fall von zwei Reihen nur die *eine* Determinante (12) und die entsprechende Unbestimmte ξ_{12}. Wegen der angenommenen Homogenität wird also: $F = c\,.\,\xi_{12}^{\varkappa}$. Aus

$$c\,.\,(12)^{\varkappa} \equiv 0 \bmod p \quad \text{[identisch in } \alpha, \beta]$$

folgt aber wegen $(12) \not\equiv 0 \bmod p$ [identisch in α, β] stets $c \equiv 0 \bmod p$: also

$$F(\xi_{ik}) \equiv 0 \bmod p \quad \text{[identisch in } \xi_{ik}],$$

womit (3) und somit (1) bewiesen ist.

Im Fall von drei Reihen gibt es die Determinanten (12), (13), (23), denen die Unbestimmten ξ_{12}, ξ_{13}, ξ_{23} entsprechen; es wird also

$$F = \sum c\,\xi_{12}^{\tau_{12}}\,\xi_{13}^{\tau_{13}}\,\xi_{23}^{\tau_{23}}.$$

Sei nun F bezw. von den Gradzahlen σ_1, σ_2, σ_3 in den Indices 1, 2, 3; dann gilt also das Gleichungssystem

$$\sigma_1 = \tau_{12} + \tau_{13}; \qquad \sigma_2 = \tau_{12} + \tau_{23}; \qquad \sigma_3 = \tau_{13} + \tau_{23}$$

mit der eindeutigen Umkehrung:

$$\tau_{12} = \tfrac{1}{2}(\sigma_1 + \sigma_2 - \sigma_3); \quad \tau_{13} = \tfrac{1}{2}(\sigma_1 - \sigma_2 + \sigma_3); \quad \tau_{23} = \tfrac{1}{2}(-\sigma_1 + \sigma_2 + \sigma_3);$$

sodaß jedes einzeln homogene F höchstens *einen* Term enthält; $F = c\,.\,\xi_{12}^{\tau_{12}}\,\xi_{13}^{\tau_{13}}\,\xi_{23}^{\tau_{13}}$, woraus wie oben das Erfülltsein von (3) und folglich (1) sich ergibt.

b) Fall von 4 Reihen α_i, β_i. Es treten die sechs Unbestimmten $\xi_{12}, \xi_{13}, \xi_{14}, \xi_{23}, \xi_{24}, \xi_{34}$ und die entsprechenden Determinanten auf; und es wird $\mathfrak{R} = (f_{12,\,34})$. Wegen

$$(4) \qquad \xi_{14}\xi_{23} \equiv \xi_{13}\xi_{24} - \xi_{12}\xi_{34} \mod \mathfrak{R}$$

ist jedes Potenzprodukt der ξ_{ik} modulo $\mathfrak{R}$ einer linearen ganzzahligen Kombination solcher Potenzprodukte kongruent, in denen das Produkt $\xi_{14}\xi_{23}$ nicht auftritt. Für jedes Polynom $F(\xi_{ik})$ gilt also:

$$(5) \qquad F(\xi_{ik}) \equiv F_0(\xi_{ik}) \mod \mathfrak{R},$$

wo F_0 das Produkt $\xi_{14}\xi_{23}$ nicht enthält; F_0 soll als das normierte Polynom der Restklasse von F bezeichnet werden. Dabei entspricht einem einzeln homogenen F, da (4) einzeln homogen ist, wieder ein einzeln homogenes F_0.

Ich setze nun:

$$(6) \qquad F_0(\xi_{ik}) = G(\xi_{ik}) + \xi_{34} \,.\, F_1(\xi_{ik}),$$

wo G kein durch ξ_{34} teilbares Potenzprodukt enthält, und weiter:

$$(7) \qquad [G]_{(\xi_{i4} = \xi_{i3})} = K(\xi_{12}, \xi_{13}, \xi_{23}).$$

Aus

$$F((ik)) \equiv 0 \mod p \text{ [identisch in } \alpha, \beta]$$

folgt nun nach (5) unter Berücksichtigung von (2):

$$(8) \qquad F_0((ik)) \equiv 0 \mod p \text{ [identisch in } \alpha, \beta];$$

und daraus durch die Spezialisierung $\alpha_4 = \alpha_3$, $\beta_4 = \beta_3$ nach (6) und (7):

$$K(ik) \equiv 0 \mod p \text{ [identisch in } \alpha, \beta].$$

Also gilt nach dem unter a) Bewiesenen:

$$K(\xi_{ik}) \equiv 0 \mod p \text{ [identisch in } \xi_{ik}],$$

und hieraus folgt, was noch unten zu zeigen ist:

$$G(\xi_{ik}) \equiv 0 \mod p \text{ [identisch in } \xi_{ik}].$$

Es wird also nach (6): $F_0(\xi_{ik}) \equiv \xi_{34} \,.\, F_1(\xi_{ik}) \mod p$ [identisch in ξ_{ik}]; und wegen $(34) \not\equiv 0 \mod p$ [identisch in α, β] folgt aus (8) auch:

$$F_1((ik)) \equiv 0 \mod p \text{ [identisch in } \alpha, \beta];$$

wo $F_1(\xi_{ik})$ wieder normiert und von niedrigerem Grad in ξ_{34} ist. Die endlich oftmalige Wiederholung des Verfahrens führt somit zu:

$$F_0(\xi_{ik}) \equiv \xi_{34}^{\lambda} F_\lambda(\xi_{ik}) \mod p,$$

wo F_λ nullten Grades in ξ_{34} ist, also nach dem eben Bewiesenen modulo p verschwindet. Somit ergibt sich:

$$(9) \qquad \begin{aligned} F_0^{3}(\xi_{ik}) &\equiv 0 \bmod p \text{ [identisch in } \xi_{ik}\text{], oder} \\ F(\xi_{ik}) &\equiv 0 \bmod (\mathfrak{N}, p) \text{ [identisch in } \xi_{ik}\text{],} \end{aligned}$$

womit die Umkehrung von (2) und damit (1) bewiesen ist; zugleich in der schärferen Aussage für $F_0(\xi_{ik})$[1]).

Es ist noch nachzutragen, daß aus $K(\xi_{ik}) \equiv 0 \bmod p$ auch $G(\xi_{ik}) \equiv 0 \bmod p$ folgt; d. h. aus $G(\xi_{ik}) \not\equiv 0 \bmod p$ auch $K(\xi_{ik}) \not\equiv 0 \bmod p$; nur zu diesem Teil des Induktionsschlusses wird die Normierung gebraucht. Da durch die Substitution $\xi_{i4} = \xi_{i3}$ die einzelnen Potenzprodukte von G nicht verschwinden, könnte ein Verschwinden von K nur dadurch eintreten, daß verschiedenen Potenzprodukten in G gleiche in K entsprechen; und das kann wieder nur dann der Fall sein, wenn diese Potenzprodukte sich nur um eine Vertauschung der Indices 3 und 4 unterscheiden. Sei nun ein solches, nach Voraussetzung normiertes, Potenzprodukt etwa gegeben durch $\xi_{12}^{\tau_{12}} \xi_{13}^{\tau_{13}} \xi_{14}^{\tau_{14}} \xi_{24}^{\tau_{24}}$. Wegen der vorausgesetzten Einzelhomogenität in jedem Index muß einem Ersetzen von 3 durch 4 in einem ξ_{ik} ein Ersetzen von 4 durch 3 in einem andern entsprechen; ein inbetrachtkommendes Glied würde also statt $\xi_{13}\xi_{24}$ das Produkt $\xi_{14}\xi_{23}$ enthalten, was durch die Normierung ausgeschlossen ist. Ganz entsprechend könnte gegen $\xi_{12}^{\tau_{12}} \xi_{13}^{\tau_{13}} \xi_{14}^{\tau_{14}} \xi_{24}^{\tau_{24}}$ wieder nur ein Glied sich aufheben, das den Faktor $\xi_{14}\xi_{23}$ enthält, was durch die Normierung ausgeschlossen ist. Somit folgt aus $K(\xi_{ik}) \equiv 0 \bmod p$ auch $G(\xi_{ik}) \equiv 0 \bmod p$.

c) Der Fall von n Reihen α_i, β_i läßt sich nun durch volle Induktion erledigen, in genauer Verallgemeinerung des eben behandelten Falles von vier Reihen[2]). Ich zeige zuerst wieder, daß für jedes Polynom F eine Kongruenz statthat:

$$(10) \qquad F(\xi_{ik}) \equiv F_0(\xi_{ik}) \bmod \mathfrak{N} \quad \text{[identisch in } \xi_{ik}\text{],}$$

wo also $\mathfrak{N} = (f_{ik,\,rs})$ und wo $F_0(\xi_{ik})$ normiert ist. Die Normierung soll dabei in folgendem bestehen: Sind i, k, r, s vier verschiedene Indices, und $i < k < r < s$, so soll F_0 das Produkt $\xi_{is}\xi_{kr}$ nicht enthalten; d. h. es sollen die Indices des einen ξ nicht beide zwischen denen des andern liegen. Diese Normierung stellt noch keine Einschränkung dar für den Fall von zwei und drei Reihen, wo also jedes Polynom normiert

1) Diese schärfere Aussage bedeutet, daß jede Restklasse nur ein einziges normiertes Polynom enthält.

2) Der Beweis gilt auch noch für $n = 3$.

ist; auch im Fall von vier Reihen existiert, wie gezeigt, ein normiertes Polynom für jede Restklasse. Die Existenz eines normierten Polynoms für jede Restklasse kann also für $(n-1)$ Reihen als bewiesen angenommen werden.

Ein beliebiges Potenzprodukt von F sei in der Form geschrieben:

$$\xi_{i_1 n} \cdots \xi_{i_\sigma n} P,$$

wo P den Index n nicht mehr enthält. Nach Voraussetzung ist also P einem ganzzahligen normierten Polynom $P_0 \bmod \mathfrak{N}$ kongruent:

$$\xi_{i_1 n} \cdots \xi_{i_\sigma n} P \equiv \xi_{i_1 n} \cdots \xi_{i_\sigma n} P_0 \bmod \mathfrak{N}.$$

Enthält nun P_0 kein ξ_{rs} derart, daß für mindestens einen Index i dies r, s zwischen i_λ und n liegt $(i_\lambda < r < s < n)$, so ist die Normierung schon erreicht, da für das Produkt irgend zweier ξ aus P_0 die Voraussetzungen erfüllt sind. Andernfalls sei i_λ ein solcher Index, daß $i_\lambda < r < s < n$; und P_0^* das ξ_{rs} enthaltende Potenzprodukt aus P_0; wegen

$$\xi_{i_\lambda n} \xi_{rs} \equiv \xi_{i_\lambda s} \xi_{rn} - \xi_{i_\lambda r} \xi_{sn} \bmod \mathfrak{N}$$

kommt dann:

$$\begin{aligned} &\xi_{i_1 n} \cdots \xi_{i_\lambda n} \cdots \xi_{i_\sigma n} P_0^* \\ &\equiv \xi_{i_1 n} \cdots \xi_{rn} \cdots \xi_{i_\sigma n} Q + \xi_{i_1 n} \cdots \xi_{sn} \cdots \xi_{i_\sigma n} R \bmod \mathfrak{N}, \\ &\equiv \xi_{i_1 n} \cdots \xi_{rn} \cdots \xi_{i_\sigma n} Q_0 + \xi_{i_1 n} \cdots \xi_{sn} \cdots \xi_{i_\sigma n} R_0 \bmod \mathfrak{N}, \end{aligned}$$

wo Q_0 und R_0 wieder die normierten Polynome der Restklassen von Q und R bedeuten; Q_0 und R_0 existieren nach Voraussetzung, da Q und R den Index n nicht enthalten, und sind sicher ganzzahlige Polynome, da Q und R einfach Potenzprodukte sind. Wegen $i_\lambda < r < s < n$ folgt aber auch:

$$((n-i_1) + \cdots (n-r) + \cdots (n-i_\sigma)) < ((n-i_1) + \cdots (n-i_\lambda) + \cdots (n-i_\sigma)),$$
$$((n-i_1) + \cdots (n-s) + \cdots (n-i_\sigma)) < ((n-i_1) + \cdots (n-i_\lambda) + \cdots (n-i_\sigma)).$$

Durch den Übergang von P_0^* zu Q_0, R_0 hat also diese Differenzensumme abgenommen; und da diese Summe notwendig $\geqq \sigma$ ist, muß das Verfahren für jedes Potenzprodukt P_0^* nach endlich vielen Schritten abbrechen. Somit kommt:

$$F(\xi_{ik}) = \sum \xi_{i_1 n} \cdots \xi_{i_\sigma n} \cdot P^{(i)} \equiv \sum \xi_{j_1 n} \cdots \xi_{j_\sigma n} \cdot S_0^{(j)} = F_0(\xi_{ik}) \bmod \mathfrak{N},$$

wo das ganzzahlige Polynom $F_0(\xi_{ik})$ notwendig normiert ist, da sich sonst die Differenzensumme noch weiter erniedrigen ließe; damit ist (10) bewiesen. Da weiter nach Voraussetzung bei $(n-1)$

Indices jedem einzeln homogenen F ein einzeln homogenes normiertes F_0 entspricht, zeigt das angegebene Verfahren, daß dies auch für n Indices der Fall ist; es handelt sich daher im folgenden nur um einzeln homogene Polynome.

Ich setze nun wieder, analog wie unter b):

(11) $$F_0(\xi_{ik}) = G(\xi_{ik}) + \xi_{n-1,n} F_1(\xi_{ik})\ ^{1)},$$

wo G kein durch $\xi_{n-1,n}$ teilbares Potenzprodukt enthält, und weiter:

(12) $$[G]_{(\xi_{in} = \xi_{i,n-1})} = K(\xi_{ik});$$

wie die Definition der Normierung zeigt, ist neben G auch K normiert.

Dann besteht wieder zwischen G und K ein-eindeutiges Entsprechen; denn es können nur solchen verschiedenen Potenzprodukten aus G gleiche in K entsprechen, die sich nur um eine Vertauschung der Indices n und $(n-1)$ unterscheiden. Sei nun etwa G vom Grade ϱ im Index $(n-1)$, vom Grade σ im Index n, und seien $\varkappa_1 \leqq \varkappa_2 \cdots \leqq \varkappa_\varrho \leqq \varkappa_{\varrho+1} \cdots \leqq \varkappa_{\varrho+\sigma}$ die in einem Potenzprodukt von G mit $(n-1)$ und n verbundenen Indices, deren Anzahl $\varrho+\sigma$ sein muß, da nach (11) die Unbestimmte $\xi_{n-1,n}$ in G nicht auftritt. Dann ist wegen der Normierung das Potenzprodukt von der Form:

$$\xi_{\varkappa_1, n-1} \cdots \xi_{\varkappa_\varrho, n-1}\, \xi_{\varkappa_{\varrho+1}, n} \cdots \xi_{\varkappa_{\varrho+\sigma}, n} \cdot P(\xi_{ik}),$$

wo $P(\xi_{ik})$ die Indices n und $(n-1)$ nicht mehr enthält. Jede Vertauschung von n mit $(n-1)$ gibt für $i \neq j$ den Faktor: $\xi_{in}\xi_{j,n-1}$ mit $i < j$; so daß also $j, n-1$ zwischen i und n liegt, das Glied nicht normiert wäre. G enthält somit zu festen $\varkappa_1, \varkappa_\varrho, \ldots, \varkappa_{r+\sigma}$ und festem P nur ein Potenzprodukt, womit gezeigt ist, daß verschiedenen Potenzprodukten aus G verschiedene in K entsprechen.

(13) Aus $K(\xi_{ik}) \equiv 0 \bmod p$ folgt somit $G(\xi_{ik}) \equiv 0 \bmod p$.

---

1) Die (10) und (11) entsprechende Identität:

$$F((ik)) = F_0((ik)) = G((ik)) + (n-1, n) \cdot F_1((ik))$$

stellt ein Analogon zu der Clebsch-Gordanschen Reihenentwicklung dar, sobald man die beiden Reihen $(n-1)$ und (n) durch Variabelnreihen x, y, also $(n-1, n)$ durch (xy) ersetzt. Während es sich aber bei Clebsch-Gordan um Entwicklung nach Polaren handelt, entspricht die Normierung von F_0 „Anfangsgliedern" von Polaren, und dadurch läßt sich die bei Clebsch-Gordan nicht geltende Ganzzahligkeit erzwingen. Der bekannte Beweis, daß unter Vernachlässigung der Ganzzahligkeit die $f_{ik, rs}$ eine Basis von $\mathfrak{M}$ bilden, wird gerade vermöge der Clebsch-Gordanschen Reihenentwicklung geführt (Gordan-Kerschensteiner: Vorlesungen über Invariantentheorie S. 132).

Aus

$$F((ik)) \equiv 0 \bmod p \text{ [identisch in } \alpha, \beta]$$

folgt nun nach (10) unter Berücksichtigung von (2):

(14) $$F_0((ik)) \equiv 0 \bmod p \text{ [identisch in } \alpha, \beta]$$

und daraus durch die Spezialisierung: $\alpha_n = \alpha_{n-1}$, $\beta_n = \beta_{n-1}$ nach (11) und (12):

(15) $$K((ik)) \equiv 0 \bmod p \text{ [identisch in } \alpha, \beta].$$

Da $K(\xi_{ik})$ normiert ist und nur $(n-1)$ Indices enthält, kann nach dem unter a) und unter b) (9) Bewiesenen wegen (15) vorausgesetzt werden:

$$K(\xi_{ik}) \equiv 0 \bmod p \text{ [identisch in } \xi_{ik}]$$

und daraus folgt nach (13):

$$G(\xi_{ik}) \equiv 0 \bmod p \text{ [identisch in } \xi_{ik}].$$

Nach (11) kommt also:

$$F_0(\xi_{ik}) \equiv \xi_{n-1,n} F_1(\xi_{ik}) \bmod p \text{ [identisch in } \xi_{ik}],$$

und wegen $(n, n-1) \not\equiv 0 \bmod p$ [identisch in α, β] folgt aus (14) auch:

$$F_1((ik)) \equiv 0 \bmod p \text{ [identisch in } \alpha, \beta];$$

wo $F_1(\xi_{ik})$ wieder normiert und von niedrigerem Grad in $\xi_{n-1,n}$ als F_0 ist. Die endlich oftmalige Wiederholung des Verfahrens führt somit zu:

$$F_0(\xi_{ik}) \equiv \xi_{n-1,n}^{\lambda} F_\lambda(\xi_{ik}) \bmod p,$$

wo F_λ nullten Grades in $\xi_{n-1,n}$ ist, also nach dem eben Bewiesenen modulo p verschwindet. Somit ergibt sich:

(16) $$\begin{aligned} &F_0(\xi_{ik}) \equiv 0 \bmod p \text{ [identisch in } \xi_{ik}] \text{ oder} \\ &F\,(\xi_{ik}) \equiv 0 \bmod (\mathfrak{R}, p) \text{ [identisch in } \xi_{ik}] \end{aligned}$$

Damit ist aber die Umkehrung von (2) und folglich (1) bewiesen; und zugleich für $F_0(\xi_{ik})$ die beim Induktionsschluß benutzte schärfere Aussage auch im Falle von n Indices, also allgemein-giltig, bewiesen.

§ 4. Die Endlichkeit der ganzzahligen Invarianten endlicher Gruppen.

Der in meiner Note „der Endlichkeitssatz der Invarianten endlicher Gruppen“ [1]) gegebene elementare Endlichkeitsbeweis läßt vermöge der speziellen Endlichkeitssätze III und IV des § 1 die Verschärfung auf Ganzzahligkeit zu; während der gewöhnliche, auf den Hilbertschen Satz von der Modulbasis sich stützende Beweis diese Ganzzahligkeit nicht liefert. Denn auch die Heranziehung der ganzzahligen Modulbasis gibt nur eine Darstellung, bei der eine beliebig hohe Potenz der ganzen Zahl h — der Ordnung der Gruppe — in den Nenner tritt.

Ich schließe mich durchweg den Bezeichnungen der erwähnten Note an. Die endliche Gruppe $\mathfrak{H}$ bestehe aus den h linearen Transformationen (von nicht verschwindender Determinante) [2]) $A_1 \dots A_h$, wobei durch A_k die lineare Transformation:

$$x_1^{(k)} = \sum_{\nu=1}^{n} a_{1\nu}^{(k)} x_\nu, \cdots x_n^{(k)} = \sum_{\nu=1}^{n} a_{n\nu}^{(k)} x_\nu$$

oder abkürzend $(x^{(k)}) = A_k(x)$ dargestellt sei. Die Gruppe $\mathfrak{H}$ führt also die Reihe (x) mit den Elementen $x_1, \dots x_n$ über in die Reihen $(x^{(k)})$ mit den Elementen $x_1^{(k)}, \dots x_n^{(k)}$. Dabei seien die Koeffizienten $a_{i\nu}^{(k)}$ als ganze algebraische Zahlen vorausgesetzt; das bedeutet insofern keine Beschränkung der Allgemeinheit, als nach einem Satz von J. Schur [3]) jede endliche Gruppe linearer Transformationen einer solchen ähnlich ist, für die die $a_{i\nu}^{(k)}$ ganze algebraische Zahlen werden.

Unter einer ganzzahligen (absoluten) Invariante der Gruppe sei eine solche ganze rationale Funktion von $x_1, \dots x_n$ verstanden, mit ganzen Zahlen des aus den $a_{i\nu}^{(k)}$ abgeleiteten algebraischen Zahlkörpers $\mathfrak{K}$ als Koeffizienten,

---

1) Math. Ann. 77, S. 89 (1915).

2) Es würde genügen, vorauszusetzen, daß $\mathfrak{H}$ einer abstrakten endlichen Gruppe einstufig isomorph ist; d. h. wenn eine Transformationsdeterminante der A verschwindet, soll $\mathfrak{H}$ einer Gruppe $\begin{pmatrix}\mathfrak{G} & 0\\ 0 & 0\end{pmatrix}$ ähnlich sein, wo die Determinanten der Substitutionen von $\mathfrak{G}$ nicht verschwinden. Die Invarianten von $\mathfrak{H}$ decken sich aber in diesem Fall mit den Invarianten von $\mathfrak{G}$, sodaß die Bedingung des Nichtverschwindens der Determinanten tatsächlich keine Beschränkung der Allgemeinheit vorstellt.

3) Über Gruppen linearer Substitutionen mit Koeffizienten aus einem algebraischen Zahlkörper. Satz VII. Math. Ann. 71. S. 355 (1912).

die bei Anwendung von $A_1 \ldots A_h$ identisch ungeändert bleibt; für eine solche Invariante $f(x)$ gilt also:

$$(1) \qquad f(x) = f(x^{(1)}) = \cdots = f(x^{(h)}) = \frac{1}{h} \cdot \sum_{k=1}^{h} f(x^{(k)}).$$

Umgekehrt wird auch jede ganzzahlige symmetrische Funktion der h Größenreihen $(x^{(k)})$ eine ganzzahlige Invariante der Gruppe.

Es bezeichne nun $\omega_1, \omega_2 \ldots \omega_\varrho$ eine Basis aller ganzen Zahlen aus $\mathfrak{K}$, so daß jede Invariante sich darstellen läßt als

$$f(x) = \omega_1 f_1(x) + \cdots + \omega_\varrho f_\varrho(x),$$

wo $f_1(x), \ldots f_\varrho(x)$ ganze rationale Zahlkoeffizienten besitzen. Nach (1) kommt somit:

$$h.f(x) = \omega_1 \sum_{k=1}^{h} f_1(x^{(k)}) + \cdots + \omega_\varrho \sum_{k=1}^{h} f_\varrho(x^{(k)}).$$

Hier sind die einzelnen Summen symmetrische Funktionen der h Größenreihen $(x^{(k)})$, und zwar mit ganzen rationalen Zahlkoeffizienten; also nach Satz III ganz und rational, mit ganzen rationalen Zahlkoeffizienten, durch endlich viele ganzrationalzahlige symmetrische Funktionen dieser Reihen ausdrückbar. Diese Basisfunktionen $I_1(x), \ldots I_\sigma(x)$ werden dabei nach dem obigen ganzzahlige Invarianten der Gruppe, mit im allgemeinen *ganzen algebraischen* Zahlkoeffizienten, da $\mathfrak{H}$ neben A nicht alle algebraisch-konjugierten Transformationen enthalten wird.

Es gilt somit eine Darstellung

$$(2) \qquad h \cdot f(x) = \omega_1 G_1(I) + \cdots + \omega_\varrho G_\varrho(I),$$

wo $G_1, \ldots G_\varrho$ ganze rationale Zahlkoeffizienten besitzen.

Sind nun $w_1, \ldots w_\varrho$ Unbestimmte, und durchläuft

$$L = w_1 G_1(I) + \cdots + w_\varrho G_\varrho(I)$$

alle Linearformen der w, denen durch die Substitution $w_i = \frac{\omega_i}{h}$ Invarianten $f(x)$ in der Darstellung (2) entsprechen, so ergibt sich nach dem Hilbertschen Satz von der ganzzahligen Modulbasis

$$(3) \qquad L = A_1(I) L_1 + \cdots + A_\tau(I) L_\tau,$$

wo die A_i ganze rationale Zahlkoeffizienten besitzen und wo wegen der Linearität aller L in den w die A_i die w nicht mehr enthalten.

Entsprechen nun durch die Substitution: $\omega_i = \frac{w_i}{h}$ den $L_1 \ldots L_\tau$

die Invariauten $\mathfrak{J}_1(x) \ldots \mathfrak{J}_\tau(x)$, so geht (3) durch diese Substitution über in:

$$(4) \qquad f(x) = A_1(I) \cdot \mathfrak{J}_1(x) + \cdots + A_\tau(I)\,\mathfrak{J}_\tau(x).$$

Diese Darstellung (4) zeigt, daß $I_1(x), \ldots I_\sigma(x)$; $\mathfrak{J}_1(x), \ldots \mathfrak{J}_\tau(x)$ eine Basis aller ganzzahligen Invarianten der Gruppe $\mathfrak{H}$ bilden, derart, daß jede Invariante sich ganz und rational, mit ganzen rationalen Zahlkoeffizienten, durch diese endlich vielen darstellen läßt.

Hat die Gruppe $\mathfrak{H}$ insbesondere die Eigenschaft, daß sie neben jeder Transformation A auch alle A', A'' ... mit algebraisch-konjugierten Zahlkoeffizienten enthält, so werden die ganzrationalzahligen symmetrischen Funktionen der h Reihen $(x^{(k)})$ Funktionen von x mit ganzen rationalen Zahlkoeffizienten; dies gilt insbesondere auch von den Basisfunktionen der symmetrischen Funktionen. Beschränkt man sich also auf Invarianten $f(x)$ mit ganzen rationalen Zahlkoeffizienten, so besitzen auch die Basisfunktionen $I_1(x), \ldots I_\sigma(x)$; $\mathfrak{J}_1(x) \ldots \mathfrak{J}_\tau(x)$ ganze rationale Zahlkoeffizienten. Es gilt also unter dieser speziellen, im allgemeinen für die Gruppen $\mathfrak{H}$ nicht erfüllten Voraussetzung, die ganzzahlige Endlichkeit auch für den Teilbereich der ganzrationalzahligen Invarianten.

Über die Umwandlung des Cholesterins in Cholancarbonsäure.

Von

A. Windaus.

Vorgelegt in der Sitzung vom 16. Mai 1919.

Die Untersuchungen der letzten Jahre haben bewiesen, daß Cholesterin und Cholsäure zu den hydroaromatischen Verbindungen gehören und beide vier hydrierte Ringe enthalten. Cholesterin ist ein einwertiger, einfach ungesättigter sekundärer Alkohol von der Formel $C_{27}H_{46}O$, Cholsäure ist eine gesättigte Trioxymonocarbonsäure von der Formel $C_{24}H_{40}O_5$, sie enthält also 3 Atome Kohlenstoff weniger als Cholesterin.

Daß diese beiden Verbindungen, die im tierischen Organismus die einzigen Vertreter einer sehr eigenartigen Stoffklasse darstellen, nahe chemische und genetische Beziehungen zu einander besitzen, ist schon seit langer Zeit vermutet worden. Eine Stütze für die Annahme kann man in dem Umstande sehen, daß die Cholsäure und einige ihrer Derivate ähnliche Farbenreaktionen geben wie Cholesterin. Indessen sind solche aus Farbenreaktionen abgeleiteten Schlußfolgerungen immer etwas unsicher, und es fehlt daher bislang der Beweis, daß die Cholsäure ein durch Oxydation gebildetes Abbauprodukt des Cholesterins ist.

Diesen Beweis zu erbringen habe ich mich in jahrelang fortgesetzten Experimenten bemüht und bin endlich auch zum Ziel gelangt.

Als Ausgangsmaterial für meine Versuche diente das Cholestan, der gesättigte Stammkohlenwasserstoff des Cholesterins, der entsteht, wenn man im Cholesterin die Doppelbildung reduziert

und die Hydroxylgruppe durch Wasserstoff ersetzt. Wird Cholestan mit Chromsäureanhydrid energisch oxydiert, zerfällt es in Aceton und in eine schön krystallisierte Monocarbonsäure $C_{24}H_{40}O_2$ nach folgender Gleichung:

$$(CH_3)_2CH\,CH_2\,C_{23}H_{39} + 4O = (CH_3)_2CO + HOOC - C_{23}H_{39} + H_2O.$$

Eine Säure von derselben Formel $C_{24}H_{40}O_2$ haben nun Wieland und Weil[1]) vor einigen Jahren bereitet, indem sie die drei Hydroxylgruppen der Cholsäure durch Wasserstoff ersetzen. Diese von den genannten Forschern als Cholancarbonsäure bezeichnete Säure stimmt mit der Säure aus Cholestan nahe überein, sodaß ich zunächst die Identität der beiden für wahrscheinlich hielt. Durch das Entgegenkommen des Herrn Prof. Wieland bin ich in den Besitz seiner Cholancarbonsäure gelangt und habe mich durch einen sorgfältigen Vergleich überzeugt, daß die beiden Säuren zwar außerordentlich ähnlich, aber nicht identisch sind; vor allem ist hervorzuheben, daß eine Mischung der nahe bei einander schmelzenden Säuren eine wesentliche Schmelzpunktserniedrigung zeigt .Augenscheinlich handelt es sich um zwei nahe verwandte Isomere. Solchen sehr ähnlichen Isomeren ist man schon mehrmals in der Cholesterinreihe begegnet; so kennt man neben dem normalen Cholestan ein durch Umlagerung gebildetes Diastereomeres, das Pseudocholestan; und besonders bemerkenswert ist es in diesem Zusammenhange, daß sich das natürlich vorkommende bakterielle Reduktionsprodukt des Cholesterins, das Koprosterin, nicht vom Cholestan, sondern vom Pseudocholestan ableitet. Es erschien darum nicht ausgeschlossen, daß auch die Cholancarbonsäure ein Derivat des Pseudocholestans sei.

Um diese Möglichkeit zu prüfen, habe ich mir das Pseudocholestan in größerer Menge bereitet und es ebenfalls der Oxydation mit Chromsäureanhydrid unterworfen. Der Reaktionsverlauf ist hier ganz derselbe, es entsteht Aceton und eine Monocarbonsäure $C_{24}H_{40}O_2$, die ein Isomeres der Säure aus Cholestan darstellt. Diese neue Säure $C_{24}H_{40}O_2$ ist nun tatsächlich mit der Cholancarbonsäure von Wieland und Weil identisch, sie zeigt denselben Schmelzpunkt und Mischschmelzpunkt und stimmt auch in allen anderen Eigenschaften mit ihr überein. Besonders geeignet zum Nachweis der Identität ist der Äthylester, der in langen Nadeln krystallisiert und bei 91—92° schmilzt.

Nach diesen Ermittlungen besteht also kein Zweifel mehr an

1) Ztschr. f. physiol. Ch. **80**, 287 [1912].

dem nahen genetischen Zusammenhang zwischen Cholesterin und Cholsäure, wie er durch folgende Formel wiedergegeben wird

$$C_{27}H_{46}O + 5O = CH_3COCH_3 + C_{24}H_{40}O_5.$$

---

10 g Cholestan wurden in 400 ccm siedendem Eisessig gelöst, auf 80° abgekühlt und allmählich mit 20 g Chromsäureanhydrid versetzt; die Lösung wurde dann 12 Stunden auf 80° unter ständigem Rühren erhitzt, darauf Natriumbisulfit und verdünnte Schwefelsäure hinzugefügt und die Mischung mit Äther ausgeschüttelt. Aus der ätherischen Lösung wurden in der üblichen Weise ein neutraler und ein saurer Anteil isoliert; der neutrale Anteil besteht nur aus unverändertem Cholestan; der saure Anteil wurde in wenig 90 proz. Essigsäure gelöst und einige Tage stehen gelassen; die innerhalb dieser Zeit abgeschiedenen Krystalle (etwa 1.2 g) wurden abgesaugt und nochmals aus Essigsäure umkystallisiert. Zur vollständigen Reinigung der Säure eignet sich besonders die Überführung in das schwer lösliche Kalium- oder Natriumsalz. Zu diesem Zweck wurden die Krystalle in Äther gelöst und die ätherische Lösung mit 10 proz. Kalilauge durchgeschüttelt, das abgeschiedene Kaliumsalz abgesaugt, mit 10 proz. Kalilauge gewaschen, darauf in heißem Eisessig gelöst und nunmehr die Lösung zur Wiedergewinnung der Säure mit verdünnter Salzsäure gefällt. Die so gereinigte und erneut umkrystallisierte Säure schmilzt bei 162°, mit Wielands Cholancarbonsäure vom Schmelzpunkt 160° gemischt beginnt sie unter 150° zu sintern, ist bei 152° trübflüssig und gibt bei etwa 155—156° eine klare Schmelze.

Analyse: 4.219 mg Sbst.: 12.34 mg CO_2, 4.27 mg H_2O.

$C_{24}H_{40}O_2$	Ber.	C 79.93 %		H 11.19 %
	Gef.	„ 79.77 „		„ 11.32 „.

Der Methylester der Säure, in der üblichen Weise aus Methylalkohol und ein wenig konz. Schwefelsäure bereitet, bildet Nadeln, die bei 79—80° schmelzen.

4.114 mg Sbst.: 12.065 mg CO_2, 4.16 mg H_2O.

$C_{25}H_{42}O_2$	Ber.	C 80.14	H 11.31
	Gef.	„ 79.98	„ 11.32.

Der Äthylester krystallisiert in Blättchen und schmilzt bei 79°, während der Äthylester der Cholancarbonsäure erst bei 91—92° schmilzt.

Die Oxydation des Pseudocholestans wurde genau ebenso vorgenommen wie diejenige des Cholestans. Auch die Verarbeitung der Reaktionsprodukte war genau dieselbe.

Die aus dem sauren Anteil isolierte Säure schmolz bei 159 bis 160° und gab mit der Cholancarbonsäure keine Schmelzpunktserniedrigung; sie krystallisierte wie diese in Spießen und war in den meisten organischen Lösungsmitteln sehr schwer löslich.

Analyse: 2.780 mg Sbst.: 8.170 mg CO_2, 2.768 mg H_2O.

$C_{24}H_{40}O_2$	Ber. C 79.93	H 11.19
	Gef. „ 80.17	„ 11.14.

Der Äthylester der Säure krystallisierte unmittelbar aus der Lösung aus, als die Säure in wenig kochendem absol. Alkohol gelöst und mit einigen Tropfen konz. Schwefelsäure versetzt wurde. Nach dem Umkrystallisieren aus 80 proz. Alkohol schmolz der Ester bei 91—92° und krystallisierte in prachtvollen langen Nadeln. Der zum Vergleich bereitete Ester der Cholancarbonsäure verhielt sich genau ebenso und gab mit dem eben erwähnten keine Schmelzpunktserniedrigung.

Zur Variationsrechnung.

Erste Mitteilung.

Von

Leon Lichtenstein in Berlin.

Vorgelegt von E. Landau in der Sitzung vom 21. Februar 1919.

Variationsprobleme sind besonders geartete Randwertaufgaben. Bei vorgeschiebenen Grenz- oder Randbedingungen sind hier Funktionen zu bestimmen, die vorgegebenen Integralen extremale Werte erteilen. Es liegt darum nahe, sich bei Behandlung der Variationsprobleme derjenigen Methoden zu bedienen, die in der Theorie der Randwertaufgaben ausgebildet worden sind. Dieser Weg ist bei zweidimensionalen Variationsproblemen von vornherein der gegebene, wenn es sich um die Konstruktion des Feldes und die hinreichenden Bedingungen handelt[1]. Aber auch bei eindimensionalen Aufgaben werden sich dabei, wie die folgenden Ausführungen zeigen, Vorteile ergeben.

Die vorliegende Arbeit beschäftigt sich mit dem einfachsten Problem bei festem und bei freiem Rand.

Es sei $f(x, y, t)$ eine Funktion, die für alle x in dem abgeschlossenen Intervalle $a \leqq x \leqq b$, kürzer in $[ab]$, für alle (x, y) in einem einfach zusammenhängenden Gebiete $\mathfrak{R}$ und alle reellen t analytisch und regulär ist. Betrachten wir das Integral

$$(1) \qquad \int_a^b f(x, y, y')\, dx.$$

1) Vgl. L. Lichtenstein, Untersuchungen über zweidimensionale reguläre Variationsprobleme. I. Das einfachste Problem bei fester Begrenzung. Jacobische Bedingung und die Existenz des Feldes. Verzweigung der Extremalflächen. Monatshefte für Mathematik und Physik 28 (1917), S. 3—51.

Es sei $\overset{0}{y}$ eine für $a \leqq x \leqq b$ erklärte analytische und reguläre Funktion von x, die so beschaffen ist, daß die Kurve $y = \overset{0}{y}(x)$ der Eulerschen Differentialgleichung

$$f_{y'y'}\frac{d^2y}{dx^2} + f_{y'y}\frac{dy}{dx} + f_{y'x} - f_y = 0 \tag{2}$$

genügt und ganz im Innern des Gebietes $\mathfrak{K}$ enthalten ist. Wir nehmen an, daß auf $y = \overset{0}{y}(x)$ die Legendresche Bedingung in der schärferen Form

$$f_{y'y'} > 0 \tag{3}$$

erfüllt ist. Betrachten wir ferner die Jacobische Differentialgleichung

$$\begin{gathered}\frac{d}{dx}\left(p\frac{du}{dx}\right) + ku = 0, \\ p = f_{y'y'}(x, \overset{0}{y}(x), \overset{0}{y}'(x)), \\ k = -f_{yy}(x, \overset{0}{y}(x), \overset{0}{y}'(x)) + \frac{d}{dx}f_{y'y}(x, \overset{0}{y}'(x), \overset{0}{y}(x)).\end{gathered} \tag{4}$$

Die dritte für das Eintreten eines Minimums notwendige Bedingung wird in der Regel in der folgenden Form angegeben:

Die Jacobische Differentialgleichung (4) darf keine partikuläre Lösung haben, die in zwei Punkten ξ_1 und ξ_2 des Intervalles $[ab]$, von denen wenigstens einer in seinem Innern liegt, verschwindet [1]).

Hat die Gleichung (4) ein partikuläres Integral, das in a und b verschwindet, für $a < x < b$ jedoch von Null verschieden ist, so kann die zweite Variation gleich Null gemacht werden. Im allgemeinen liegt auch dann kein Minimum vor [2]).

In einer vor zwei Jahren erschienenen Arbeit über die zweidimensionalen Variationsprobleme habe ich die Jacobische Bedingung auf eine andere Form gebracht [3]). Sie lautet auf den eindimensionalen Fall übertragen wie folgt.

Damit ein Minimum möglich sei, muß der zu dem ersten Randwertproblem gehörige kleinste positive Eigenwert der aus der Jacobischen Differentialgleichung (4) durch Hinzufügung eines Parameters gewonnenen Differentialgleichung

---

1) Vgl. z. B. O. Bolza, Vorlesungen über Variationsrechnung. Leipzig und Berlin 1909, S. 54—87.

2) Vgl. O. Bolza, a. a. O.

3) Vgl. L. Lichtenstein, a. a. O. S. 17—18.

(5) $$\frac{d}{dx}\left(p\frac{du}{dx}\right)+\lambda k u = 0$$

mindestens gleich 1 sein.

Es seien λ_j die Eigenwerte der Differentialgleichung (5), φ_j die zugehörigen den Beziehungen

$$\int_a^b k\varphi_j\varphi_l\,dx = 0 \quad (j \neq l), \qquad \int_a^b k\varphi_j^2\,dx = \frac{\lambda_j}{|\lambda_j|}$$

gemäß normierten, für $x = a$ und $x = b$ verschwindenden Eigenfunktionen. Ist in $[ab]$

$$k \leqq 0, \qquad \int_a^b k dx \neq 0,$$

so sind alle Eigenwerte negativ. Die Jacobische Bedingung ist stets erfüllt. In diesem Falle wird man den kleinsten positiven Eigenwert gleich $+\infty$ setzen.

Der vorstehende Satz folgt fast unmittelbar aus der in § 2 des ersten Kapitels abgeleiteten neuen Formel für die zweite Variation (Formel (23)). Man gewinnt sie, wenn man die variierte Funktion in der Form

(6) $$\overset{0}{y} + \varepsilon u = \overset{0}{y} + \varepsilon \sum \frac{\lambda_j}{|\lambda_j|}\varphi_j x_j, \qquad x_j = \int_a^b k\varphi_j u dx$$

schreibt und damit die Variationsaufgabe gewissermaßen in ein Minimumproblem einer Funktion von unendlich vielen Variablen x_j umwandelt[1]). Dieser Gedanke erweist sich namentlich dann als besonders nützlich, wenn der kleinste positive Eigenwert λ_1 gerade gleich 1 ist. Es gelingt hier unter Heranziehung höherer Variationen eine beliebig weit fortsetzbare Reihe explizit angebbarer Kriterien aufzustellen, die zu entscheiden gestatten, ob ein (schwaches) Minimum eintritt oder nicht[2]).

1) Man vergleiche D. Hilbert, Wesen und Ziele einer Analysis der unendlich vielen unabhängigen Variabeln, Rendiconti del Circolo Matematico di Palermo 27 (1909), S. 59—74, wo sich S. 66 ein dahinzielender Hinweis findet.

2) Ganz anders geartete Kriterien zur Entscheidung dieser Frage sind von den Herren Kneser, Osgood und Lindeberg angegeben worden. Vgl. A. Kneser, Zur Variationsrechnung, Math. Annalen 50 (1898), S. 27—50; Lehrbuch der Variationsrechnung, Braunschweig 1900, S. 93—97; W. F. Osgood, On the existence of a minimum of the integral $\int_{x_0}^{1} F(x,y,y')\,dx$, when x_0 und x_1 are conjugate points, and the geodetics on an ellipsoid of revolution. A revision of a theorem of Kneser's, Transactions of the American Mathematical Society 2 (1901), S. 166—182; E. Lindeberg, Zur Theorie der Maxima und Minima ein-

Ist $\lambda_1 > 1$, so läßt sich die Extremale $\overset{0}{y}$ in ein Feld von Extremalen einbetten. Die Bestimmung eines Feldes läßt sich ohne weiteres auf eine Randwertaufgabe zurückführen. Diese wird in § 4 des ersten Kapitels in einer sehr einfachen Weise durch successive Approximationen erledigt.

Ist $\lambda_1 = 1$, so liegt, wie ich in dem zweidimensionalen Falle schon früher gezeigt habe, Verzweigung der Extremalkurven vor[1]).

Das zweite Kapitel behandelt das einfachste Problem bei freiem Rand. Es ist als ein Vorteil der in dieser Arbeit benutzten Methode anzusehen, daß die Jacobische Bedingung in einer ganz ähnlichen Gestalt erscheint wie bei festem Rand. Es muß wieder der kleinste positive Eigenwert λ_1 einer gewissen erweiterten Jacobischen Differentialgleichung, diesmal unter Zugrundelegung der Randbedingungen

$$(7)\quad \begin{gathered} p(a)\frac{du}{dx} + \lambda h_1 u = 0 \text{ für } x = a, \quad p(b)\frac{du}{dx} - \lambda h_2 u = 0 \text{ für } x = b, \\ h_1 = f_{y'y}(a, \overset{0}{y}(a), \overset{0}{y}'(a)), \qquad h_2 = -f_{y'y}(b, \overset{0}{y}(b), \overset{0}{y}'(b)), \end{gathered}$$

mindestens gleich 1 sein. Der sonst übliche Übergang durch das Problem mit einem festen und einem freien Rand und die Betrachtung der Brennpunkte werden dabei vermieden.

Ist $\lambda_1 > 1$, so läßt sich $\overset{0}{y}$ in ein Feld von Extremalkurven einbetten. Es möge in bekannter Bezeichnungsweise die Weierstraßsche Funktion $E(x, y; p, \tilde{p})$ für alle p in einer Umgebung von $\overset{0}{p} = \frac{d\overset{0}{y}}{dx}$ und alle von p verschiedenen $\tilde{p}$ positiv sein. Es liegt dann ein starkes Minimum vor. Der besondere Fall $\lambda_1 = 1$ gestattet ganz wie bei festem Rand eine ins einzelne gehende Diskussion. Liegt hier ein Minimum vor, so ist es stets ein starkes Minimum.

Das in dieser Arbeit eingeschlagene Verfahren läßt sich, was wesentlich ist, in ähnlicher Weise bei dem einfachsten Problem der zweidimensionalen Variationsrechnung bei festem und namentlich bei freiem Rand, wo die üblichen Methoden versagen, durchführen[2]).

---

facher Integrale mit bestimmten Integrationsgrenzen, Math. Annalen 59 (1904), S. 321—331. Man vergleiche hierzu O. Bolza, Vorlesungen über Variationsrechnung, Leipzig und Berlin 1909, S. 363—364, wo auch die ältere Literatur angegeben ist. Die vorstehenden Arbeiten beschränken sich übrigens nicht auf die Betrachtung des schwachen Minimums.

1) Vgl. die in der Fußnote 1 S. 161 zitierte Arbeit, S. 35—51.

2) Vgl. die vor kurzem in der Math. Zeitschrift Bd. 5 (1919) erschienene zweite Abhandlung der Serie „Untersuchungen über zweidimensionale Variationsprobleme“.

Das gleiche Verfahren dürfte sich ferner auch auf manche schwierigere Aufgaben der Variationsrechnung übertragen lassen. Auch die hier gegebene Form der Jacobischen Bedingung dürfte sich dort, in geeigneter Weise modifiziert, wiederfinden.

Nur noch einige Bemerkungen methodischen Inhalts. Die Entwicklungen des ersten Kapitels (feste Grenzen) stützen sich auf den Satz von der Existenz der Eigenwerte der Differentialgleichung (5) und die in § 1 wiedergegebenen Reihenentwicklungen der Integrale $\int_a^b ku^2\,dx$, $\int_a^b p\left(\frac{du}{dx}\right)^2 dx$ und der am Rande verschwindenden Greenschen Funktion der Differentialgleichung $\frac{d}{dx}\left(p\frac{du}{dx}\right) = 0$. Diese Sätze habe ich auf den ersten Seiten meiner in der Anmerkung 1 S. 166 unter a) genannten Arbeit abgeleitet. Als einzige Hilfsmittel erscheinen dort der Hilbertsche Fundamentalsatz über die orthogonale Transformation vollstetiger quadratischer Formen mit unendlich vielen Unbekannten und ein bekannter Satz des Herrn Toeplitz über beschränkte Formen.

Ganz analoge Bemerkungen gelten für den Inhalt des zweiten Kapitels (freie Grenzen).

Erstes Kapitel.

Feste Grenzen.

§ 1.

Betrachten wir die für $x = a$ und $x = b$ verschwindenden Eigenfunktionen der Differentialgleichung

$$\frac{d}{dx}\left(p\frac{du}{dx}\right) + \lambda k u = 0. \tag{5}$$

Nimmt k in $[ab]$ positive und negative Werte an, so gibt es unendlich viele positive und unendlich viele negative Eigenwerte. Die positiven Eigenwerte mögen $\lambda_1 < \lambda_2 < \lambda_3 < \ldots$, die negativen $\lambda_{-1} > \lambda_{-2} > \ldots$ heißen. Die zugehörigen den Beziehungen

$$\int_a^b k\varphi_j\varphi_l\,dx = 0 \quad (j \neq l), \qquad \int_a^b k\varphi_j^2\,dx = \frac{\lambda_j}{|\lambda_j|} \tag{8}$$

gemäß normierten Eigenfunktionen werden mit

(9) $$\varphi_1, \varphi_2, \ldots; \quad \varphi_{-1}, \varphi_{-2}, \ldots$$

bezeichnet.

Nach bekannten Sätzen gilt für jede in $[ab]$ stetige, für $x = a$ und $x = b$ verschwindende Funktion, die in $[ab]$ abteilungsweise stetige Ableitung hat,

(10) $$\int_a^b ku^2\,dx = \sum \frac{\lambda_j}{|\lambda_j|}\left(\int_a^b k\varphi_j u\,dx\right)^2 = \sum \frac{\lambda_j}{|\lambda_j|} x_j^2.$$

Es gilt ferner die weitere Formel

(11) $$\int_a^b p\left(\frac{du}{dx}\right)^2 dx = \sum |\lambda_j|\left(\int_a^b k\varphi_j u\,dx\right)^2 = \sum |\lambda_j| x_j^2 \text{ [1]}.$$

Die unendliche Reihe

(12) $$\sum \frac{\varphi_j(x)\,\varphi_j(y)}{|\lambda_j|}$$

konvergiert für alle x und y in dem Rechtecke $a \leqq x \leqq b$, $a \leqq y \leqq b$ unbedingt und gleichmäßig [2]. Es ist ferner

(13) $$\sum \frac{\varphi_j(x)\,\varphi_j(y)}{|\lambda_j|} = G(x, y),$$

unter $G(\xi, x)$ die auf dem Rande verschwindende Greensche Funktion der Differentialgleichung $\frac{d}{dx}\left(p\,\frac{du}{dx}\right) = 0$ verstanden [3].

§ 2.

Es sei u irgendeine für $a \leqq x \leqq b$ stetige, für $x = a$ und $x = b$ verschwindende Funktion, die in $[ab]$ abteilungsweise stetige Ableitung hat und zugleich einer Ungleichheit

(14) $$\left|\frac{du}{dx}\right| \leqq M \quad (M \text{ konstant})$$

---

1) Vgl. L. Lichtenstein, a) Zur Analysis der unendlich vielen Variabeln. I. Entwicklungssätze der Theorie gewöhnlicher linearer Differentialgleichungen zweiter Ordnung, Rendiconti del Circolo Matematico di Palermo 38 (1914), S. 113—166 [S. 126]; b) Zweite Abhandlung, Mathematische Zeitschrift 3 (1919), S. 127—160 [S. 128].

2) Mehr als dies, die aus den absoluten Beträgen einzelner Glieder gebildete Reihe konvergiert gleichmäßig.

3) Vgl. L. Lichtenstein, a. a. O. a) S. 128.

genügt. Aus $u(x) = \int_a^x \frac{du}{dx}\,dx$ folgt einerseits

(15) $$|u| < M(b-a),$$

andererseits in bekannter Weise

(16) $$u^2 = \left(\int_a^x \frac{du}{dx}\,dx\right)^2 \leqq \int_a^x \left(\frac{du}{dx}\right)^2 dx \int_a^x dx \leqq (b-a)\int_a^b \left(\frac{du}{dx}\right)^2 dx.$$

Es sei weiter ε_0 eine dem absoluten Betrage nach so kleine Größe, daß die Kurve

(17) $$y = \overset{0}{y} + \varepsilon u$$

für $a \leqq x \leqq b$ und für alle $|\varepsilon| \leqq \varepsilon_0$ ganz im Innern des Gebietes $\mathfrak{R}$ liegt. Geht $\overset{0}{y}$ in $\overset{0}{y} + \varepsilon u$ über, so erhält das Integral

(18) $$J = \int_a^b f(x, \overset{0}{y}, \overset{0}{y}')\,dx,$$

da seine erste Variation verschwindet, die Zunahme

(19) $$\varDelta J = \varepsilon^2 J_2 + \varepsilon^3 J_3 + \cdots.$$

Wir wollen von nun an der Einfachheit halber für

(20) $$f_{y'y'}(x, \overset{0}{y}, \overset{0}{y}'), \qquad f_{y'y'y'}(x, \overset{0}{y}, \overset{0}{y}'), \ldots$$

kürzer

(21) $$\overset{0}{f}_{y'y'}, \qquad \overset{0}{f}_{y'y'y'}, \ldots$$

schreiben. Die zweite Variation läßt sich durch teilweise Integration auf die Form

(22) $$\begin{aligned} J_2 &= \frac{1}{2}\int_a^b \left[\overset{0}{f}_{y'y'} \left(\frac{du}{dx}\right)^2 + 2\overset{0}{f}_{y'y}\, u\,\frac{du}{dx} + \overset{0}{f}_{yy}\, u^2\right] dx \\ &= \frac{1}{2}\int_a^b \left\{f_{y'y'}\left(\frac{du}{dx}\right)^2 + \left(\overset{0}{f}_{yy} - \frac{d}{dx}\overset{0}{f}_{y'y}\right) u^2\right\} dx \\ &= \frac{1}{2}\int_a^b \left[p\left(\frac{du}{dx}\right)^2 - ku^2\right] dx \end{aligned}$$

bringen. Wegen

(10) $$\int_a^b ku^2\,dx = \sum \frac{\lambda_j}{|\lambda_j|}\, x_j^2, \qquad x_j = \int_a^b ku\varphi_j\,dx$$

und

(11) $$\int_a^b p\left(\frac{du}{dx}\right)^2 dx = \sum |\lambda_j| x_j^2$$

gilt

(23) $$\boxed{J_2 = \frac{1}{2}\sum |\lambda_j|\left(1-\frac{1}{\lambda_j}\right) x_j^2.}$$

Aus dieser Formel folgt fast unmittelbar.

Damit die zweite Variation $J_2 \geqq$ *sei, ist notwendig und hinreichend, daß der kleinste positive Eigenwert*

$$\lambda_1 \geqq 1$$

sei.

Andererseits ist

(24) $$\varDelta J - \varepsilon^2 J_2 = \frac{\varepsilon^3}{3}\int_a^b \left[\tilde{f}_{y'y'y'}\left(\frac{du}{dx}\right)^3 + 3\tilde{f}_{y'y'y}\left(\frac{du}{dx}\right)^2 u + 3\tilde{f}_{y'yy}\frac{du}{dx}u^2 + \tilde{f}_{yyy}u^3\right] dx,$$

unter $f_{y'y'y}$ u. s. w. zur Abkürzung die Ausdrücke

(25) $$f_{y'y'y'} = \tilde{f}_{y'y'y}[x, \overset{0}{y} + \theta\varepsilon u, \quad \overset{0}{y}' + \theta\varepsilon u'], \qquad 0 < \theta < 1$$

. . . .

verstanden. Wegen (14) ist für $|\varepsilon| \leqq \varepsilon_0$

(26) $$\left|\int_a^b \tilde{f}_{y'y'y'}\left(\frac{du}{dx}\right)^3 dx\right| \leqq \alpha_1 M \int_a^b \left(\frac{du}{dx}\right)^2 dx,$$

unter α_1 wie später unter $\alpha_2, \alpha_3, \ldots$ eine positive Konstante verstanden. Wir finden ferner wegen (15)

(27) $$\left|\int_a^b \tilde{f}_{y'y'y}u\left(\frac{du}{dx}\right)^2 dx\right| \leqq \alpha_2 M \int_a^b \left(\frac{du}{dx}\right)^2 dx$$

und mit Rücksicht auf (14) und (16)

(28) $$\left|\int_a^b \tilde{f}_{y'yy}u^2\frac{du}{dx}dx\right| \leqq \alpha_3 M \int_a^b u^2 dx \leqq \alpha_4 M \int_a^b \left(\frac{du}{dx}\right)^2 dx,$$

(29) $$\left|\int_a^b \tilde{f}_{yyy}u^3 dx\right| \leqq \text{Max.}\ |\tilde{f}_{yyy}|\ \text{Max.}\ |u| \int_a^b u^2 dx \leqq \alpha_5 M \int_a^b \left(\frac{du}{dx}\right)^2 dx.$$

Demnach ist

$$(30)\qquad |\varDelta J - \varepsilon^2 J_2| \leqq \varepsilon^3 \alpha_6 M \int_a^b \left(\frac{du}{dx}\right)^2 dx \leqq \varepsilon^3 \alpha_7 M \int_a^b p \left(\frac{du}{dx}\right)^2 dx.$$

1. Ist $\lambda_1 < 1$, so wird wegen (23) und (8) für $u = \varphi_1$ gewiß $J_2 < 0$, demnach wegen (30) für alle hinreichend kleinen Werte von $|\varepsilon|$

$$(31)\qquad \varDelta J < 0.$$

Es liegt kein Minimum vor.

Damit ein Minimum eintreten kann, muß demnach der kleinste positive Eigenwert der durch die Einführung eines Parameters erweiterten Jacobischen Differentialgleichung (5), $\lambda_1 \geqq 1$ sein.

2. Ist $\lambda_1 > 1$, so ist offenbar $J_2 > 0$. Es gilt ferner, wie man sich leicht überzeugt,

$$(32)\qquad J_2 \geqq \frac{1}{2}\left(1 - \frac{1}{\lambda_1}\right) \sum |\lambda_j| x_j^2,$$

demnach

$$(33)\qquad J_2 \geqq \frac{1}{2}\left(1 - \frac{1}{\lambda_1}\right) \int_a^b p \left(\frac{du}{dx}\right)^2 dx.$$

Das Gleichheitszeichen gilt nur für die Funktionen $u = \text{const.}\ \varphi_1$ und $u = 0$.

Aus (30) und (33) folgt jetzt für alle der Beziehung (14) genügenden $\frac{du}{dx}$ und alle dem absoluten Betrage nach hinreichend kleinen Werte von ε,

$$(34)\qquad |\varepsilon| \leqq \varepsilon_1 \leqq \varepsilon_0:$$

$$(35)\qquad \varDelta J \geqq \varepsilon^2 \left(\frac{1}{2}\left(1 - \frac{1}{\lambda_1}\right) - \varepsilon \alpha_7 M\right) \int_a^b p \left(\frac{du}{dx}\right)^2 dx,\quad \frac{1}{2}\left(1 - \frac{1}{\lambda_1}\right) - \varepsilon \alpha_7 M > 0.$$

Das Gleichheitszeichen gilt nur für $u \equiv 0$.

Ist in $[ab] \ldots k \leqq 0$, $\int_a^b k\,dx \neq 0$, so sind alle Eigenwerte negativ.

Alsdann ist

$$(36)\qquad J_2 \geqq \frac{1}{2} \int_a^b p \left(\frac{du}{dx}\right)^2 dx,$$

$$(37)\qquad \varDelta J \geqq \varepsilon^2 \left(\frac{1}{2} - \alpha \varepsilon_7 M\right) \int_a^b p \left(\frac{du}{dx}\right)^2 dx,\quad \frac{1}{2} - \varepsilon \alpha_7 M > 0.$$

Die Ungleichheit (37) folgt aus (35) für $\lambda_1 = +\infty$. Den besonderen Fall durchweg negativer Eigenwerte kann man in den Fall $\lambda_1 > 1$ einordnen, indem man $\lambda_1 = +\infty$ setzt.

Aus (35) folgt zunächst:

Ist der kleinste positive Eigenwert der erweiterten Jacobischen Differentialgleichung (5), $\lambda_1 > 1$, so liegt gewiß ein schwaches Minimum vor.

3. Ist schließlich $\lambda_1 = 1$, so kann J_2 verschwinden. Dies tritt nach (23) dann und nur dann ein, wenn $u = \text{const.}\, \varphi_1$ ist. Ob und unter welchen Bedingungen dann ein schwaches Minimum eintreten kann, wird in § 5 näher untersucht werden.

§ 3.

Es ist nicht schwer zu zeigen, daß die in dem vorhergehenden Paragraphen abgeleitete notwendige Bedingung der Jacobischen Bedingung in der üblichen Fassung völlig gleichwertig ist. Nach dieser darf es, damit ein Minimum eintreten kann, keine Lösung der Jacobischen Differentialgleichung (4) geben, die in zwei Punkten von $[ab]$ verschwindet, von denen wenigstens einer im Innern dieses Intervalles liegt[1]).

Es sei $\lambda_1 \geqq 1$, und es sei $\bar{u}(x)$ eine in $[ab]$ reguläre Lösung der Differentialgleichung (4), die etwa in a und in dem Punkte ξ $(a < \xi < b)$ verschwindet. Wir setzen

$$(37) \qquad J_2 = \frac{1}{2}\int_a^{\xi}\left(p\left(\frac{du}{dx}\right)^2 - ku^2\right)dx + \frac{1}{2}\int_{\xi}^{b}\left(p\left(\frac{du}{dx}\right)^2 - ku^2\right)dx.$$

Bezeichnen wir die zu dem Intervalle $[a\xi]$ gehörigen Eigenfunktionen und Eigenwerte der Differentialgleichung (5) mit $\bar{\varphi}_j(x)$ und $\bar{\lambda}_j$. Es gilt die zu (23) analoge Formel

$$(38) \qquad \int_a^{\xi}\left(p\left(\frac{d\bar{u}}{dx}\right)^2 - k\bar{u}^2\right)dx = \sum |\bar{\lambda}_j|\left(1 - \frac{1}{\bar{\lambda}_j}\right)\left(\int_a^{\xi} k\bar{u}\bar{\varphi}_j\, dx\right)^2.$$

Offenbar ist nun $\bar{u}(x)$ eine zu dem Intervalle $[a\xi]$ und dem Eigenwert 1 gehörige Eigenfunktion der Differentialgleichung (5). Nach § 2 ist in (37) das erste Integral rechter Hand für $u = \bar{u}$ gleich Null. Setzt man jetzt $u = \bar{u}$ in $[a\xi]$, $u = 0$ in $[\xi b]$, so ist wegen (37)

---

1) Vgl. z. B. O. Bolza, Vorlesungen über Variationsrechnung, Leipzig und Berlin 1909, S. 54—87.

(39) $$J_2 = \frac{1}{2}\int_a^b \left(p\left(\frac{du}{dx}\right)^2 - ku^2\right)dx = 0,$$

was dem in § 2 bewiesenen Resultat widerspricht.

Es ist jetzt leicht zu sehen, daß die zu dem kleinsten positiven Eigenwert λ_1 gehörige Eigenfunktion $\varphi_1(x)$ im Innern von $[ab]$ nicht verschwindet[1]). Es genügt hierzu die Differentialgleichung

(40) $$\frac{d}{dx}\left(p\frac{du}{dx}\right) + \lambda k_1 u = 0, \qquad k_1 = \lambda_1 k$$

zu betrachten. Ihr kleinster positiver Eigenwert ist offenbar gleich 1. Die zugehörige Eigenfunktion, d. h. die Funktion φ_1 kann den soeben durchgeführten Betrachtungen gemäß in $[ab]$ nicht verschwinden.

§ 4.

Es sei $\lambda_1 > 1$. Wir beweisen, daß die Extremale $\overset{0}{y}$ sich in ein Feld von Extremalen einbetten läßt.

Es sei ε eine dem absoluten Betrage nach hinreichend kleine Zahl und $y = \overset{0}{y} + \varepsilon z$ eine zu $\overset{0}{y}$ benachbarte Extremale. Es sei

(41) $$y(a) = \overset{0}{y}(a) + \varepsilon A, \quad y(b) = \overset{0}{y}(b) + \varepsilon B.$$

Wegen $\lambda_1 > 1$ ist das erste Randwertproblem der Differentialgleichung (4) stets und zwar nur in einer Weise lösbar. Es sei $\chi(x)$ diejenige in $[ab]$ reguläre Lösung der Differentialgleichung (4), die den Randbedingungen $\chi(a) = A$, $\chi(b) = B$ genügt. Wir setzen

(42) $$y = \overset{0}{y} + \varepsilon\chi + \varepsilon z$$

und erhalten aus der Eulerschen Differentialgleichung (2) zur Bestimmung der für $x = a$ und $x = b$ verschwindenden Funktion z eine Differentialgleichung von der Form

(43) $$\frac{d}{dx}\left(p\frac{dz}{dx}\right) + kz = \varepsilon F(z, z', z''),$$

unter F eine für alle $(x, \overset{0}{y} + \varepsilon\chi + \varepsilon z)$ in $\mathfrak{R}$ und für alle reellen z' und z'' analytische und reguläre Funktion von $\varepsilon, x, z, z', z''$ verstanden.

Bezeichnet $\Pi(x)$ irgendeine in $[ab]$ stetige Funktion und ist $\mathfrak{G}(\xi, x)$ die Greensche Funktion von (4), so ist nach bekannten Sätzen

1) Dies gilt gleichgültig, ob $\lambda_1 >, =$ oder < 1 ist.

$$Z(\xi) = -\int_a^b \mathfrak{G}(\xi, x)\, \Pi(x)\, dx \tag{44}$$

die in $[ab]$ nebst ihrer ersten und zweiten Ableitung stetige, kürzer „reguläre", für $x = a$ und $x = b$ verschwindende Lösung der Gleichung

$$\frac{d}{dx}\left(p\frac{dZ}{dx}\right) + kZ = \Pi\,[1]. \tag{45}$$

Ist $|\Pi| < N$, so folgt aus (44) und (45)

$$|Z|, \left|\frac{dZ}{dx}\right|, \left|\frac{d^2Z}{dx^2}\right| < \beta N \qquad (\beta \text{ konstant}). \tag{46}$$

Aus (43) folgt nun

$$z(\xi) = -\varepsilon \int_a^b \mathfrak{G}(\xi, x)\, F(z(x), z'(x), z''(x))\, dx. \tag{47}$$

Es möge für alle $|\varepsilon| < \varepsilon^*$, alle x in $[ab]$ und alle $|z|, |z'|, |z''| < \delta^*$

$$|F|, |F_z|, |F_{z'}|, |F_{z''}| < N \tag{48}$$

sein [2]. Die Integro-Differentialgleichung (47) wird durch successive Approximationen gelöst. Wir setzen

$$z_1(\xi) = -\varepsilon \int_a^b \mathfrak{G}(\xi, x)\, F(0, 0, 0)\, dx,$$

$$\cdots\cdots\cdots\cdots\cdots \tag{49}$$

$$z_\nu(\xi) = -\varepsilon \int_a^b \mathfrak{G}(\xi, x)\, F(z_{\nu-1}(x), z'_{\nu-1}(x), z''_{\nu-1}(x))\, dx \quad (\nu = 2, 3, \ldots)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

Für hinreichend kleine $|\varepsilon|$, etwa $|\varepsilon| < \varepsilon^{(0)} \leqq \varepsilon^*$ ist sicher

$$|z_1|, |z'_1|, |z''_1| < \delta^*. \tag{50}$$

Aus (48) und (46) folgt

$$|z_2|, |z'_2|, |z''_2| < \beta\varepsilon N, \tag{51}$$

somit für

---

1) Vgl. z. B. D. Hilbert, Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. Zweite Mitteilung, Gött. Nachrichten 1904, S. 213—259 [S. 220—221].

2) Selbstverständlich muß die Kurve $\overset{0}{y} + \varepsilon\chi + \varepsilon z$ für alle $|\varepsilon| < \varepsilon^*$ und $|z| < \delta^*$ in $\mathfrak{R}$ liegen.

(52) $$|\varepsilon| < \varepsilon^{(1)} = \text{Min.}\left(\varepsilon^{(0)}, \frac{\delta^*}{\beta N}\right)$$

gewiß

(53) $$|z_2|, |z_2'|, |z_2''| < \delta^*.$$

Offenbar gilt für alle ν, sofern $|\varepsilon| < \varepsilon^{(1)}$ angenommen wird,

(54) $$|z_\nu|, |z_\nu'|, |z_\nu''| < \delta^*.$$

Es sei jetzt zur Abkürzung

(55) $$\zeta_\nu = z_{\nu+1} - z_\nu \quad (\nu = 1, 2, \ldots)$$

gesetzt. Aus (49) ergibt sich

(56) $$\zeta_\nu = -\varepsilon \int_a^b \mathfrak{G}(\xi, x)\,[\tilde{F}_z \zeta_{\nu-1} + \tilde{F}_{z'} \zeta_{\nu-1}' + \tilde{F}_{z''} \zeta_{\nu-1}''] \, dx,$$

$$\tilde{F}_z = F_z(z_{\nu-1} + \theta\zeta_{\nu-1}, z_{\nu-1}' + \theta\zeta_{\nu-1}', z_{\nu-1}'' + \theta\zeta_{\nu-1}''), \ldots \quad 0 < \theta < 1.$$

Wegen (50) und (53) ist

(57) $$|\zeta_1|, |\zeta_1'|, |\zeta_1''| < 2\delta^*,$$

demnach wegen (56), (48) und (46)

(58) $$|\zeta_2|, |\zeta_2'|, |\zeta_2''| < 6N\delta^*\varepsilon\beta.$$

Für

(59) $$|\varepsilon| < \varepsilon^{(2)} = \text{Min.}\left(\varepsilon^{(1)}, \frac{1}{6N\beta}\right)$$

folgt aus (58)

(60) $$|\zeta_2|, |\zeta_2'|, |\zeta_2''| < \delta^*$$

und, wie man durch eine Wiederholung dieser Überlegung sieht, allgemein

(61) $$|\zeta_\nu|, |\zeta_\nu'|, |\zeta_\nu''| < \frac{\delta^*}{2^{\nu-1}} \qquad (\nu = 1, 2, \ldots).$$

Die unendlichen Reihen

(62) $$z_1 + \sum \zeta_\nu, \quad z_1' + \sum \zeta_\nu', \quad z_1'' + \sum \zeta_\nu''$$

konvergieren für alle x in $[ab]$ und alle $|\varepsilon| \leqq \varepsilon^{(2)}$ unbedingt und gleichmäßig. Die Funktion $z = z_1 + \sum \zeta_\nu$ ist die gesuchte Lösung der Differentialgleichung (43). Die Funktion $y = \overset{0}{y} + \varepsilon\chi + \varepsilon z$ stellt eine Schar von Extremalen dar. Ist speziell $A = \overline{A} > 0$, $B = \overline{A} > 0$,

so ist wegen $\lambda_1 > 1$ für alle x in $[ab] \ldots \chi > 0$[1]). Aus (42) und (47) folgt

$$\left[\frac{\partial y}{\partial \varepsilon}\right]_{\varepsilon=0} = \chi > 0. \tag{63}$$

Die Schar y stellt ein Feld von Extremalen dar.

Wie wir vorhin bewiesen haben, läßt sich ein Wert $\bar{\delta}$ angeben, so daß, wie man auch die Werte y_a und y_b den Beziehungen $|y_a - \overset{0}{y}(a)| < \bar{\delta}$, $|y_b - \overset{0}{y}(b)| < \bar{\delta}$ gemäß wählt, es stets eine Extremale y gibt, die die Randbedingungen $y(a) = y_a$, $y(b) = y_b$ erfüllt[2]). Für $y_a \to \overset{0}{y}(a)$, $y_b \to \overset{0}{y}(b)$ ist, wie sich leicht zeigen läßt, gleichmäßig

$$y \to \overset{0}{y},\ y' \to \overset{0}{y}{}',\ y'' \to \overset{0}{y}{}''. \tag{64}$$

Ist $\bar{\lambda}_1$ der kleinste positive Eigenwert, der zu der Extremale y gehört, so ist ferner gleichmäßig $\bar{\lambda}_1 \to \lambda_1$. Ist ϑ irgendein positiver Wert $< \lambda_1 - 1$, so kann man eine Zahl $\bar{\bar{\delta}} \leqq \bar{\delta}$ angeben, sodaß für alle

$$|y_a - \overset{0}{y}(a)| \leqq \bar{\bar{\delta}},\quad |y_b - \overset{0}{y}(b)| \leqq \bar{\bar{\delta}} \tag{65}$$

$$\bar{\lambda}_1 > 1 + \vartheta \tag{66}$$

wird. Jede den Beziehungen (65) genügende Extremale läßt sich darum in ein Feld von Extremalen einbetten. Man kann ohne Schwierigkeiten zeigen, daß es sich für alle Extremalen y der Schar (65) eine Zahl $\varepsilon^{(3)} \leqq \varepsilon^{(2)}$ angeben läßt, so daß die zu den Randwerten $y_a + \varepsilon \bar{A}$, $y_b + \varepsilon \bar{A}$ für alle $|\varepsilon| \leqq \varepsilon^{(3)}$ gehörenden Extremalen ein y umgebendes Feld bilden. Man kann $\bar{\bar{\delta}}$ so klein wählen, daß in dem von diesem Felde bestrichenen Flächenstück (für jedes Paar Randwerte y_a, y_b in (65)) alle Extremalen der Schar (65) enthalten sind. Diese Bemerkung wird für uns später in § 2 des zweiten Kapitels von Nutzen sein.

Ist in allen Punkten (x, y) eines Feldes von Extremalen um $\overset{0}{y}$ und alle von p verschiedenen $\tilde{p}$ die Weierstraßsche Funktion $E(x, y;\ p(x, y), \tilde{p}) > 0$, so liegt, wie man weiß, ein starkes Minimum vor. Als Vergleichsfunktionen können alle stetigen Funktionen $y(x)$ herangezogen werden, die abteilungsweise stetige Ableitung

1) Die Funktion χ kann in $[ab]$ nirgends gleich Null sein, weil es sonst im Gegensatz zu den Ergebnissen des § 3 eine reguläre Lösung geben würde, die in mindestens zwei im Innern von $[ab]$ gelegenen Punkten verschwindet.

2) Es ist übrigens nicht schwer zu zeigen, daß es auch nur eine Extremale gibt, die diesen Bedingungen genügt.

haben, für $x = a$ und $x = b$ mit $\overset{\circ}{y}$ übereinstimmen und überdies so beschaffen sind, daß die Kurve $x, y(x)$ das betrachtete Feld um $\overset{\circ}{y}$ nicht verläßt[1]).

§ 5.

Wir wenden uns jetzt einer eingehenden Betrachtung des Falles $\lambda_1 = 1$ zu.

Wir setzen

(67) $$u = x_1 \varphi_1 + v$$

und erhalten wegen

(68) $$\int_a^b \left[p\left(\frac{d\varphi_1}{dx}\right)^2 - k\varphi_1^2\right] dx = 0, \qquad \int_a^b \left[p\frac{d\varphi_1}{dx}\frac{dv}{dx} - k\varphi_1 v\right] dx = 0,$$

(69) $$\begin{aligned} J_2 &= \frac{1}{2}\int_a^b \left[p\left(\frac{dv}{dx}\right)^2 - kv^2\right] dx \\ &= \frac{1}{2}\sum{}' |\lambda_j|\, x_j^2 - \frac{1}{2}\sum{}' \frac{\lambda_j}{|\lambda_j|} x_j^2 = \frac{1}{2}\sum{}' |\lambda_j| \left(1 - \frac{1}{\lambda_j}\right) x_j^2. \end{aligned}$$

Durch das Zeichen $\sum'$ soll angedeutet werden, daß bei der Summation der Wert $j = 1$ außer Betracht bleibt.

Es ist

(70) $$\int_a^b p\left(\frac{du}{dx}\right)^2 dx = \sum |\lambda_j|\, x_j^2 = x_1^2 + \int_a^b p\left(\frac{dv}{dx}\right)^2 dx,$$

(71) $$\int_a^b p\left(\frac{dv}{dx}\right)^2 dx \leqq \int_a^b p\left(\frac{du}{dx}\right)^2 dx,$$

(72) $$x_1^2 \leqq \int_a^b p\left(\frac{du}{dx}\right)^2 dx \leqq (b-a)\,\text{Max.}\, p\, M^2.$$

Es gilt ferner

(73) $$v = \int_a^x \frac{dv}{dx}\,dx, \quad v^2 \leqq \int_a^b dx \int_a^b \left(\frac{dv}{dx}\right)^2 dx \leqq \frac{b-a}{\text{Min.}\, p}\int_a^b p\left(\frac{du}{dx}\right)^2 dx.$$

Wir finden schließlich

(74) $$\left|\frac{dv}{dx}\right| \leqq \left|\frac{du}{dx}\right| + |x_1|\left|\frac{d\varphi_1}{dx}\right| \leqq M + (b-a)^{\frac{1}{2}}\,\text{Max.}\, p^{\frac{1}{2}}\, M \left|\frac{d\varphi_1}{dx}\right| \leqq \alpha_3,$$

---

1) Vgl. z. B. O. Bolza, Vorlesungen über Variationsrechnung, Leipzig und Berlin 1909, S. 121.

$$(75)\quad \begin{aligned} |v| &\leqq |u| + |x_1|\,|\varphi_1| < M(b-a) + |x_1|\,|\varphi_1| \\ &\leqq M(b-a) + (b-a)^{\frac{1}{2}}\,\mathrm{Max.}\,p^{\frac{1}{2}} M\,|\varphi_1| < \alpha_9. \end{aligned}$$

Betrachten wir jetzt die dritte Variation

$$(76)\qquad J_3 = \frac{1}{3!}\int_a^b \left[\overset{0}{f}_{y'y'y'}\left(\frac{du}{dx}\right)^3 + \cdots + f_{yyy}u^3\right]dx.$$

Setzt man hier $u = x_1\varphi_1 + v$, so erhält man unter dem Integralzeichen eine ganze rationale Funktion dritten Grades von v und $\frac{dv}{dx}$. Die von v und $\frac{dv}{dx}$ freien Glieder sind

$$(77)\qquad x_1^3 \varDelta^3 = \frac{1}{3!}\,x_1^3 \int_a^b \left[\overset{0}{f}_{y'y'y'}\left(\frac{d\varphi_1}{dx}\right)^3 + \cdots + \overset{0}{f}_{yyy}\varphi_1^3\right]dx.$$

In bekannter Weise findet man als eine für das Eintreten des Minimums notwendige Bedingung die Beziehung

$$(78)\qquad \boxed{\varDelta_3 = 0.}$$

Betrachten wir jetzt die Glieder, die in bezug auf v und $\frac{dv}{dx}$ vom zweiten und dritten Grade sind. Sie sind jedes einzelne dem absoluten Betrage nach kleiner als

$$(79)\qquad \alpha_{10}\int_a^b p\left(\frac{dv}{dx}\right)^2 dx.$$

Das ist für die Glieder zweiten Grades mit Rücksicht auf (73) einleuchtend. Was die Glieder dritten Grades betrifft, so sind hier Beziehungen wie

$$(80)\quad |v^3| = |v|\,|v^2| \leqq \alpha_9 (b-a)\int_a^b \left(\frac{dv}{dx}\right)^2 dx \leqq \alpha_{11}\int_a^b \left(\frac{dv}{dx}\right)^2 dx \text{ usw.}$$

zu berücksichtigen. Wir können darum setzen

$$(81)\quad J_3 = \bar{J}_3 + x_1^2\int_a^b \left(h_1 v + h_2\frac{dv}{dx}\right)dx, \qquad |\bar{J}_3| \leqq \alpha_{12}\int_a^b p\left(\frac{dv}{dx}\right)^2 dx$$

oder, nach einer teilweisen Integration, unter $h_1, h_2, \ldots$ bekannte in $[ab]$ analytische und reguläre Funktionen verstanden,

$$(82)\qquad J_3 = \bar{J}_3 + x_1^2\int_a^b \left(h_1 - \frac{dh_2}{dx}\right) v\,dx = \bar{J}_3 + x_1^2\int_a^b hv\,dx$$

und analog

$$J_4 = \bar{J}_4 + x_1^3 \int_a^b \left(h_3 v + h_4 \frac{dv}{dx}\right) dx + x_1^4 \varDelta_4 = \bar{J}_4 + x_1^3 \int_a^b h^{(1)} v dx + x_1^4 \varDelta_4,$$

(83)

$$|\bar{J}_4| \leqq \alpha_{13} \int_a^b p \left(\frac{dv}{dx}\right)^2 dx,$$

wo $x_1^4 \varDelta_4$ die Gesamtheit der von v und $\frac{dv}{dx}$ freien Glieder in J_4 bezeichnet:

(84) $$\varDelta_4 = \frac{1}{4!} \int_a^b \left[\overset{0}{f}_{y'y'y'y'} \left(\frac{d\varphi_1}{dx}\right)^4 + \cdots + \overset{0}{f}_{yyyy} \varphi_1^4\right] dx.$$

Die Gesamtvariation des Integrals J läßt sich demnach in der Form schreiben:

$$\varDelta J = \varepsilon^2 J_2 + \varepsilon^3 J_3 + \varepsilon^4 J_4 + \varepsilon^5 R_5$$

(85) $$= \frac{\varepsilon^2}{2} \sum{}' |\lambda_j| \left(1 - \frac{1}{\lambda_j}\right) x_j^2 + \varepsilon^3 x_1^3 \int_a^b h v dx + \varepsilon^4 x_1^3 \int_a^b h^{(1)} v dx$$

$$+ \varepsilon^3 \bar{J}_3 + \varepsilon^4 \bar{J}_4 + \varepsilon^4 x_1^4 \varDelta_4 + \varepsilon^5 R_5.$$

Der Rest R_5 ist ein Integral über einer Form fünften Grades der beiden Argumente $v + x_1 \varphi_1$ und $\frac{dv}{dx} + x_1 \frac{d\varphi_1}{dx}$. Für die Gesamtheit der Glieder zweiten bis fünften Grades in bezug auf v und $\frac{dv}{dx}$ gewinnt man leicht eine Abschätzung nach oben von der Form

$$\alpha_{14} \int_a^b p \left(\frac{dv}{dx}\right)^2 dx,$$

für die übrigen Glieder eine Abschätzung von der Form

$$\alpha_{15} x_1^4,$$

sodaß man alles in allem setzen kann

(86) $$|R_5| < \alpha_{14} \int_a^b p \left(\frac{dv}{dx}\right)^2 dx + \alpha_{15} x_1^4.$$

Für (85) kann man auch schreiben

$$\Delta J = \frac{\varepsilon^2}{2}\sum{}'|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)x_j^2 + \varepsilon^3\sum{}' x_j x_1^2 \int_a^b h\varphi_j\,dx$$

$$+\varepsilon^4 \sum{}' x_j x_1^3 \int_a^b h^{(1)}\varphi_j\,dx + \varepsilon^3 \overline{J}_3 + \varepsilon^4 \overline{J}_4 + \varepsilon^4 x_1^4 \Delta_4 + \varepsilon^5 R_5$$

(87) $$= \frac{\varepsilon^2}{2}\sum{}'|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)\left\{x_j + \varepsilon x_1^2 \frac{\int_a^b h\varphi_j\,dx}{|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)} + \varepsilon^2 x_1^3 \frac{\int_a^b h^{(1)}\varphi_j\,dx}{|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}\right\}^2$$

$$+\varepsilon^3 \overline{J}_3 + \varepsilon^4 \overline{J}_4 + \varepsilon^4 x_1^4 \left\{\Delta_4 - \sum{}' \frac{\left[\int_a^b (h+\varepsilon x_1 h^{(1)})\varphi_j\,dx\right]^2}{2|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}\right\} + \varepsilon^5 R_5\,{}^{1)}.$$

Aus der Formel (87) folgt, wie wir jetzt zeigen wollen, die weitere für das Eintreten eines Minimums notwendige Bedingung

(88) $$\boxed{\Delta_4 - \sum{}' \frac{\left(\int_a^b h\varphi_j\,dx\right)^2}{2|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)} \geqq 0.}$$

Aus (88) folgt insbesondere $\Delta_4 \geqq 0$.

1. Es sei zunächst

(89) $$\Delta_4 - \sum{}' \frac{\left(\int_a^b h\varphi_j\,dx\right)^2}{2|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)} > 0.$$

---

1) Wie bald gezeigt werden wird, ist die unendliche Reihe

$$\sum{}' \varphi_j(x) \frac{\int_a^b (h+\varepsilon x_1 h^{(1)})\varphi_j\,dx}{|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}$$

gleichmäßig konvergent. Wie man leicht sieht, ist darum auch die Reihe

$$\sum{}' \frac{\left[\int_a^b (h+\varepsilon x_1 h^{(1)})\varphi_j\,dx\right]^2}{|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}$$

konvergent. Die Umformung des Textes ist demnach zulässig.

Dann ist für hinreichend kleine Werte von $\delta > 0$ und $|\varepsilon|$ gewiß

$$(90)\qquad \varDelta_4 - \sum{}' \frac{\left(\int_a^b (h + \varepsilon x_1 h^{(1)}) \varphi_j \, dx\right)^2}{2(1-\delta)\,|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} > \delta_1 > 0.$$

Der Ausdruck (87) läßt sich nun wie folgt umformen

$$\varDelta J = \frac{\varepsilon^2}{2} \sum{}' (1-\delta)\,|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right) \left\{ x_j + \varepsilon x_1^2 \frac{\int_a^b (h + \varepsilon x_1 h^{(1)}) \varphi_j \, dx}{(1-\delta)\,|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} \right\}^2$$

$$(91)\qquad + \varepsilon^4 x_1^4 \left\{ \varDelta_4 - \frac{\left(\int_a^b (h + \varepsilon x_1 h^{(1)}) \varphi_j \, dx\right)^2}{2(1-\delta)\,|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} \right\}$$

$$+ \frac{\delta \varepsilon^2}{2} \sum{}' |\lambda_j| \left(1 - \frac{1}{\lambda_j}\right) x_j^2 + \varepsilon^3 \bar{J}_3 + \varepsilon^4 \bar{J}_4 + \varepsilon^5 R_5 = \sum_{j=1}^{6} K_j.$$

Der Ausdruck K_1 ist nicht negativ, K_2 ist dem absoluten Betrage nach größer als $\delta_1 \varepsilon^4 x_1^4$. Es gilt ferner

$$(92)\qquad K_3 \geqq \left(1 - \frac{1}{\lambda_2}\right) \frac{\delta \varepsilon^2}{2} \sum{}' |\lambda_j|\, x_j^2 = \left(1 - \frac{1}{\lambda_2}\right) \frac{\delta \varepsilon^2}{2} \int_a^b p \left(\frac{dv}{dx}\right)^2 dx$$

und wegen (81, (83) und (86)

$$(93)\qquad |K_4 + K_5 + K_6| \leqq \alpha_{16} |\varepsilon|^3 \int_a^b p \left(\frac{dv}{dx}\right)^2 dx + \alpha_{17} |\varepsilon|^5 x_1^4.$$

Für alle hinreichend kleinen $|\varepsilon|$, etwa $|\varepsilon| \leqq \varepsilon_2 \leqq \varepsilon_0$, ist demnach

$$(94)\qquad |\varDelta J| \geqq \alpha_{18} \varepsilon^2 \int_a^b p \left(\frac{dv}{dx}\right)^2 dx + \alpha_{19} \varepsilon^4 x_1^4.$$

Das Gleichheitszeichen gilt nur für $x_1 = 0$, $\frac{dv}{dx} \equiv 0$, d. h. $u \equiv 0$.

Es liegt also ein schwaches Minimum vor.

2. Es möge jetzt zweitens

$$\Delta_4 - \sum' \frac{\left(\int\limits_a^b h\varphi_j dx\right)^2}{2|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)} < 0 \tag{95}$$

sein.

Dann ist für hinreichend kleine Werte von $\delta > 0$ und $|\varepsilon|$ gewiß

$$\Delta_4 - \sum' \frac{\left(\int\limits_a^b (h+\varepsilon x_1 h^{(1)})\,\varphi_j dx\right)^2}{2(1+\delta)\,|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)} < \delta_2 < 0. \tag{96}$$

Wir schreiben

$$\begin{aligned}\Delta J = \frac{\varepsilon^2}{2}\sum'(1+\delta)\,|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)&\left\{x_j + \varepsilon x_1^2 \frac{\int\limits_a^b (h+\varepsilon x_1 h^{(1)})\,\varphi_j dx}{(1+\delta)\,|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}\right\}^2\\ &+\varepsilon^4 x_1^4 \left\{\Delta_4 - \frac{\left(\int\limits_a^b (h+\varepsilon x_1 h^{(1)})\,\varphi_j dx\right)^2}{2(1+\delta)\,|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}\right\}\\ &-\delta\frac{\varepsilon^2}{2}\sum'|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)x_j^2 + \varepsilon^3 \overline{J}_3 + \varepsilon^4 \overline{J}_4 + \varepsilon^5 R_5 = \sum_{j=1}^{6} K_j'.\end{aligned} \tag{97}$$

Setzen wir einmal

$$\overline{x}_j = -\varepsilon x_1^2 \frac{\int\limits_a^b (h+\varepsilon x_1 h^{(1)})\,\varphi_j dx}{(1+\delta)\,|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}. \qquad (j \geqq 2) \tag{98}$$

Die unendliche Reihe

$$\overline{v}(x) = \sum' \overline{x}_j\,\varphi_j(x) \tag{99}$$

konvergiert, wie jetzt gezeigt werden soll, unbedingt und gleichmäßig und stellt eine in dem Intervalle $a \leqq x \leqq b$ nebst ihren Ableitungen erster und zweiter Ordnung stetige Funktion dar.

Wie in dem § 1 hervorgehoben wurde, konvergiert die unendliche Reihe

$$\sum \frac{\varphi_j(x)\,\varphi_j(y)}{|\lambda_j|} \tag{100}$$

für alle x und y in dem Rechteck $a \leqq x \leqq b$, $a \leqq y \leqq b$ unbedingt und gleichmäßig (vgl. die Fußnote 2 S. 166). Es ist ferner

$$(101) \quad \sum \frac{\varphi_j(x)\,\varphi_j(y)}{|\lambda_j|} = G(x,y) = \begin{cases} \dfrac{(b-y)\,(x-a)}{b-a} & \text{für } x \leqq y, \\[2mm] \dfrac{(b-x)\,(y-a)}{b-a} & \text{für } x \geqq y. \end{cases}$$

Offenbar konvergiert die Reihe

$$(102) \quad \sum{}' \varphi_j(x) \frac{\int\limits_a^b h^* \varphi_j\,dy}{|\lambda_j|} = \int\limits_a^b h^*(y)\,G(x,y)\,dy - \varphi_1(x) \int\limits_a^b h^* \varphi_1\,dy,$$
$$h^* = h + \varepsilon x_1 h^{(1)}$$

in $[ab]$ unbedingt und gleichmäßig und stellt eine nebst ihren Ableitungen erster und zweiter Ordnung stetige Funktion dar [1]).

Betrachten wir ferner die augenscheinlich unbedingt und gleichmäßig konvergierende Reihe

$$(103) \quad \sum{}' \varphi_j(x) \int\limits_a^b h^* \varphi_j dy \left\{ \frac{1}{|\lambda_j|\left(1-\dfrac{1}{\lambda_j}\right)} - \frac{1}{|\lambda_j|} \right\} = \sum{}' \varphi_j(x) \frac{1}{|\lambda_j|\,(\lambda_j - 1)} \int\limits_a^b h^* \varphi_j\,dy.$$

Die unendliche Reihe

$$(104) \quad \sum{}' \frac{d^2 \varphi_j(x)}{dx^2} \frac{1}{|\lambda_j|\,(\lambda_j-1)} \int\limits_a^b h^* \varphi_j dy = -\sum{}' k(x) \frac{\varphi_j(x)}{\lambda_j - 1} \frac{\lambda_j}{|\lambda_j|} \int\limits_a^b h^* \varphi_j\,dy$$

konvergiert, wie aus dem Vergleich mit der Reihe (102) hervorgeht, ebenfalls unbedingt und gleichmäßig. In bekannter Weise schließt man hieraus, daß die Funktion (103) stetige Ableitungen erster und zweiter Ordnung hat. Hierdurch ist bewiesen, daß die unendliche Reihe (99) in der Tat unbedingt und gleichmäßig konvergiert und eine nebst ihren Ableitungen erster und zweiter Ordnung stetige Funktion darstellt.

Führt man jetzt in (97) für x_1 den Wert $\bar{x}_1$, somit für v die spezielle Funktion $\bar{v}$ ein, so verschwindet K_1'. Es gilt ferner

$$(105) \quad K_2' \leqq -\delta_2 \varepsilon^4 x_1^4, \qquad K_3' \leqq -\frac{\varepsilon^2 \delta}{2}\left(1 - \frac{1}{\lambda_2}\right) \int\limits_a^b p \left(\frac{d\bar{v}}{dx}\right)^2 dx,$$

$$(106) \quad |K_4' + K_5' + K_6'| \leqq \alpha_{16} |\varepsilon|^3 \int\limits_a^b p \left(\frac{d\bar{v}}{dx}\right)^2 dx + \alpha_{17} |\varepsilon|^5 x_1^4.$$

---

1) Auch die aus den absoluten Beträgen der einzelnen Glieder gebildete Reihe konvergiert gleichmäßig.

Offenbar ist für alle hinreichend kleinen Werte $|\varepsilon|$, etwa $|\varepsilon| \leqq \varepsilon_8 \leqq \varepsilon_0$

$$\Delta J \leqq -\alpha_{18}\varepsilon^2 \int_a^b p\left(\frac{d\bar{v}}{dx}\right)^2 dx - \alpha_{19}\varepsilon^4 x_1^4. \tag{107}$$

Das Gleichheitszeichen gilt nur für $x_1 = 0$, $\bar{v} \equiv 0$.
Ein Minimum tritt nicht ein.

3. Es sei schließlich

$$\Delta_4 - \sum{}' \frac{\left(\int_a^b h\varphi_j\,dx\right)^2}{2|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} = 0. \tag{108}$$

Es bleibt dann vorläufig noch unentschieden, ob ein Minimum eintreten wird oder nicht.

Wir setzen

$$X(x) = \sum{}' \varphi_j(x) \frac{\int_a^b h\varphi_j\,dy}{|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} = \sum{}' d_j\varphi_j(x) \tag{109}$$

und

$$\begin{aligned} u(x) = x_1\varphi_1(x) + v &= x_1\varphi_1(x) - \varepsilon x_1^2 X(x) + w(x) \\ &= x_1\varphi_1(x) - \varepsilon x_1^2 X(x) + \sum{}' y_j\varphi_j(x), \end{aligned} \tag{110}$$

demnach

$$x_j = -\varepsilon x_1^2 d_j + y_j. \qquad (j \neq 1) \tag{111}$$

Es gilt

$$|X(x)| < \alpha_{20}, \qquad \left|\frac{dX(x)}{dx}\right| < \alpha_{21}, \tag{112}$$

demnach (vgl. (74) und (75))

$$|w(x)| < \alpha_{20}|\varepsilon| x_1^2 + \alpha_9, \qquad \left|\frac{dw(x)}{dx}\right| < \alpha_{21}|\varepsilon| x_1^2 + \alpha_8, \tag{113}$$

oder mit Rücksicht auf (72)

$$|w(x)|, \qquad \left|\frac{dw(x)}{dx}\right| < \alpha_{22}. \tag{114}$$

Es gilt ferner in bekannter Weise

$$|w| \leqq \alpha_{23} \int_a^b p\left(\frac{dw}{dx}\right)^2 dx. \tag{115}$$

Wir führen den Ausdruck (111) für x_j in (87) ein und erhalten mit Rücksicht auf (108) zunächst

(116)
$$\Delta J = \frac{\varepsilon^2}{2}\sum' |\lambda_j|\left(1-\frac{1}{\lambda_j}\right)\left\{y_j + \varepsilon^2 x_1^3 \frac{\int_a^b h^{(1)}\varphi_j dx}{|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}\right\}^2 + \varepsilon^3 \bar{J}_3 + \varepsilon^4 \bar{J}_4$$
$$-\varepsilon^5 x_1^5 \sum' \frac{1}{2|\lambda_j|\left(1-\frac{1}{\lambda_j}\right)}\left[2\int_a^b h\varphi_j dx \int_a^b h^{(1)}\varphi_j dx + \varepsilon x_1 \left(\int_a^b h^{(1)}\varphi_j dx\right)^2\right] + \varepsilon^5 R_5.$$

Betrachten wir jetzt den Ausdruck $\bar{J}_3$. Wir haben so die Gesamtheit der Glieder der dritten Variation J_3 bezeichnet, die nach Substitution von
$$u = x_1\varphi_1 + v$$
unter dem Integralzeichen in bezug auf v und $\frac{dv}{dx}$ vom zweiten oder dritten Grade sind. Führt man jetzt $v = -\varepsilon x_1^2 X + w$ ein, so zerfällt $\bar{J}_3$ nach dem Schema

(117)
$$\bar{J}_3 = \sum_{\nu=0}^{3} \bar{J}_{3\nu},$$

unter $\bar{J}_{3\nu}$ die Gesamtheit der Glieder verstanden, die in bezug auf w und $\frac{dw}{dx}$ vom ν-ten Grade sind. Man findet nunmehr wie vorhin für $\bar{J}_3$ ohne Mühe die Abschätzungen

(118)
$$|\bar{J}_{32}|,\ |\bar{J}_{33}| \leqq \alpha_{24}\int_a^b p\left(\frac{dw}{dx}\right)^2 dx.$$

Es ist ferner

(119)
$$\bar{J}_{31} = \varepsilon x_1^3 \sum' y_j \int_a^b (h^{(2)} + \varepsilon x_1 h^{(3)})\varphi_j dx,$$

unter $h^{(2)}$ und $h^{(3)}$, wie später unter $h^{(4)}, h^{(5)}, \ldots$ bestimmte in $[ab]$ stetige Funktionen verstanden. Aus jedem Gliede von $\bar{J}_3$, das in bezug auf v oder $\frac{dv}{dx}$ vom zweiten, darum in bezug auf $x_1\varphi_1(x)$ vom ersten Grade ist, entsteht in der Tat ein Glied von $\bar{J}_{31}$ von der Form

(120)
$$x_1\int_a^b h_5 \varepsilon x_1^2 w dx \quad \text{oder} \quad x_1\int_a^b h_6 \varepsilon x_1^2 \frac{dw}{dx} dx = -\varepsilon x_1^3 \int_a^b \frac{dh_6}{dx} w dx.$$

Aus einem Gliede von $\bar{J}_3$, das in bezug auf v oder $\frac{dv}{dx}$ vom dritten, demnach in bezug auf $x_1\varphi_1(x)$ vom nullten Grade ist, entsteht ein Glied von $\bar{J}_{31}$ von der Form

$$(121)\qquad \int\limits_a^b h_7\,\varepsilon^2 x_1^4\, w\,dx \text{ oder } \int\limits_a^b h_8\,\varepsilon^2 x_1^4 \frac{dw}{dx}\,dx = -\varepsilon^2 x_1^4 \int\limits_a^b \frac{dh_8}{dx}\, w\,dx.$$

Es gilt ferner, wie man sich in ähnlicher Weise überzeugt,

$$(122)\qquad \bar{J}_{30} = \beta_1\varepsilon^2 x_1^5 + \beta_2\varepsilon^3 x_1^6,$$

unter β_1 und β_2, wie später unter $\beta_3, \ldots$ Konstanten verstanden.

In einer ganz ähnlichen Weise läßt sich $\bar{J}_4$ umformen. Es gilt

$$(123)\qquad \begin{aligned} \bar{J}_4 &= \sum_{\nu=0}^{4} \bar{J}_{4\nu}, \\ |\bar{J}_{44}|,\ |\bar{J}_{43}|,\ |\bar{J}_{42}| &\leqq \alpha_{25} \int\limits_a^b p\left(\frac{dw}{dx}\right)^2 dx, \\ \bar{J}_{41} &= \varepsilon x_1^4 \sum{}' y_j \int\limits_a^b (h^{(4)} + \varepsilon x_1 h^{(5)} + \varepsilon^2 x_1^2 h^{(6)})\,\varphi_j\,dx, \\ \bar{J}_{40} &= \beta_3\varepsilon^2 x_1^6 + \beta_4\varepsilon^3 x_1^7 + \beta_5\varepsilon^4 x_1^8. \end{aligned}$$

Wir müssen schließlich in R_5 die fünfte und die sechste Variation gesondert betrachten. Es sei darum

$$(124)\qquad \varepsilon_5 R_5 = \varepsilon^5 J_5 + \varepsilon^6 J_6 + \varepsilon^7 R_7.$$

Wir führen in J_5 für u den Ausdruck (110) ein und setzen ähnlich wie vorhin

$$(125)\qquad J_5 = \sum_{\nu=1}^{5} J_{5\nu}.$$

Es gilt jetzt zunächst

$$(126)\qquad J_{50} = \gamma_1 x_1^5 + \gamma_2\varepsilon x_1^6 + \cdots + \gamma_6\varepsilon^5 x_1^{10},$$

unter $\gamma_1, \ldots \gamma_5$ leicht angebbare Integralausdrücke über φ_1, X, $\frac{d\varphi_1}{dx}$ und $\frac{dX}{dx}$ verstanden. Es ist weiter

$$(127)\quad J_{51} = \sum{}' y_j x_1^4 \int\limits_a^b (h^{(9)} + \varepsilon x_1 h^{(10)} + \varepsilon^2 x_1^2 h^{(11)} + \varepsilon^3 x_1^3 h^{(12)} + \varepsilon^4 x_1^4 h^{(13)})\,\varphi_j\,dx,$$

$$(128)\qquad |J_{52}|, \ldots, |J_{55}| \leqq \alpha_{26} \int\limits_a^b p\left(\frac{dw}{dx}\right)^2 dx.$$

Analoge Beziehungen gelten für J_6. Es ist nämlich

(129) $$J_6 = \sum_{\nu=1}^{6} J_{6\nu},$$

(130) $$J_{60} = \gamma_7 x_1^6 + \cdots + \gamma_{13} \varepsilon^6 x_1^{12};$$
$$J_{61} = \Sigma' y_j x_1^5 \int_a^b (h^{(14)} + \varepsilon x_1 h^{(15)} + \cdots + \varepsilon^5 x_1^5 h^{(19)}) \varphi_j dx,$$

(131) $$|J_{62}|, \ldots, |J_{66}| \leqq \alpha_{27} \int_a^b p \left(\frac{dw}{dx}\right)^2 dx.$$

Man überzeugt sich endlich leicht unter Berücksichtigung von (114) und (115), daß

(132) $$|R_7| \leqq \alpha_{28} x_1^6 + \alpha_{29} \int_a^b p \left(\frac{dw}{dx}\right)^2 dx$$

gilt. Faßt man jetzt die Entwicklungen dieses Paragraphen zusammen, so findet man, daß sich $\varDelta J$ in der Form darstellen läßt:

(133) $$\begin{aligned} \varDelta J = \frac{\varepsilon^2}{2} \Sigma' |\lambda_j| \left(1 - \frac{1}{\lambda_j}\right) \left\{ y_j + \varepsilon^2 x_1^3 \frac{\int_a^b h^{(1)} \varphi_j dx}{|\lambda_j| \left(1 - \frac{1}{\lambda_j}\right)} \right\}^2 \\ + \Sigma' \varepsilon^4 x_1^3 y_j \int_a^b \left[h^{(2)} + \varepsilon x_1 (h^{(3)} + h^{(4)} + h^{(9)}) + \varepsilon^2 x_1^2 (h^{(5)} + h^{(10)} + h^{(14)}) \right. \\ + \varepsilon^3 x_1^3 (h^{(6)} + h^{(11)} + h^{(15)}) + \varepsilon^4 x_1^4 (h^{(12)} + h^{(16)}) + \varepsilon^5 x_1^5 (h^{(13)} + h^{(17)}) + \varepsilon^6 x_1^6 h^{(18)} \\ \left. + \varepsilon^7 x_1^7 h^{(19)} \right] \varphi_j dx + (\beta_1 + \gamma_1) \varepsilon^5 x_1^5 + (\beta_2 + \beta_3 + \gamma_2 + \gamma_7) \varepsilon^6 x_1^6 + (\beta_4 + \gamma_3 + \gamma_8) \varepsilon^7 x_1^7 \\ + (\beta_5 + \gamma_4 + \gamma_9) \varepsilon^8 x_1^8 + (\gamma_5 + \gamma_{10}) \varepsilon^9 x_1^9 + (\gamma_6 + \gamma_{11}) \varepsilon^{10} x_1^{10} + \gamma_{12} \varepsilon^{11} x_1^{11} + \gamma_{13} \varepsilon^{12} x_1^{12} \\ + \bar{J}^* + \varepsilon^7 R_7, \qquad |\bar{J}^*| \leqq |\varepsilon|^3 \alpha_{30} \int_a^b p \left(\frac{dw}{dx}\right)^2 dx. \end{aligned}$$

Für (133) kann man nach einer leichten Umrechnung abgekürzt setzen

(134) $$\begin{aligned} \varDelta J = \frac{\varepsilon^2}{2} \Sigma' |\lambda_j| \left(1 - \frac{1}{\lambda_j}\right) y_j^2 + \Sigma' \varepsilon^4 x_1^3 y_j \int_a^b (l^{(1)} + \varepsilon x_1 l^{(2)} + \cdots + \varepsilon^7 x_1^7 l^{(9)}) \varphi_j dx \\ + \varkappa_1 \varepsilon^5 x_1^5 + \varkappa_2 \varepsilon^6 x_1^6 + \cdots + \varkappa_3 \varepsilon^{12} x_1^{12} + \varDelta, \\ |\varDelta| \leqq \alpha_{31} |\varepsilon|^3 \int_a^b p \left(\frac{dw}{dx}\right)^2 dx + \alpha_{32} |\varepsilon|^7 x_1^6. \end{aligned}$$

Für (134) kann man auch schreiben

$$\varDelta J = \frac{\varepsilon^2}{2} \sum{}' |\lambda_j| \left(1 - \frac{1}{\lambda_j}\right) \left\{ y_j + \varepsilon^2 x_1^3 \frac{1}{|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} \int_a^b (l^{(1)} + \varepsilon x_1 l^{(2)} + \cdots \varepsilon^7 x_1^7 l^{(8)}) \varphi_j \, dx \right\}$$

$$(135) \quad + \varkappa_1 \varepsilon^5 x_1^5 + \varepsilon^6 x_1^6 \left\{ \varkappa_2 - \frac{1}{2} \sum{}' \frac{1}{|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} \left(\int_a^b (l^{(1)} + \varepsilon x_1 (l^{(2)} + \cdots + \varepsilon^7 x_1^7 l^{(8)}) \varphi_j \, dx \right)^2 \right.$$

$$\left. + \varkappa_3 \varepsilon^7 x_1^7 + \cdots + \varkappa_8 \varepsilon^{12} x_1^{12} + \varDelta. \right.$$

Man überzeugt sich jetzt leicht, daß, damit ein Minimum vorliegen kann, jedenfalls

$$(136) \qquad \boxed{\varkappa_1 = 0}$$

sein muß. Ist diese Bedingung erfüllt und ist überdies

$$(137) \qquad \boxed{\varkappa_2 - \frac{1}{2} \sum{}' \frac{\left(\int_a^b l^{(1)} \varphi_j \, dx \right)^2}{|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} > 0,}$$

so liegt ein Minimum vor. Ist

$$\varkappa_2 - \frac{1}{2} \sum{}' \frac{\left(\int_a^b l^{(1)} \varphi_j \, dx \right)^2}{|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} < 0,$$

so tritt kein Minimum ein. Ist schließlich

$$\varkappa_2 - \frac{1}{2} \sum{}' \frac{\left(\int_a^b l^{(1)} \varphi_j \, dx \right)^2}{|\lambda_j|\left(1 - \frac{1}{\lambda_j}\right)} = 0,$$

so läßt sich eine Entscheidung aufgrund der bisherigen Betrachttungen allein nicht treffen. Die Untersuchung muß vielmehr auf Glieder höherer Ordnung ausgedehnt werden, was in ähnlicher Weise geschehen kann.

Zweites Kapitel.

Freie Grenzen.

§ 1.

Wir nehmen jetzt an, die beiden Endpunkte der Extremalkurve $y = \mathring{y}(x)$ seien auf den Geraden $x = a$, $x = b$ frei beweglich.

Es sei u irgendeine für $a \leqq x \leqq b$ stetige Funktion, die daselbst abteilungsweise stetige Ableitung hat, und den Beziehungen

$$(138) \qquad |u| \leqq M_1, \qquad \left|\frac{du}{dx}\right| \leqq M_1 \qquad (M_1 \text{ konstant})$$

genügt[1]. Es möge $\mathring{y}$ in $\mathring{y} + \varepsilon u$ übergehen. Betrachten wir die zweite Variation J_2 des Integrals

$$(1) \qquad J = \int_a^b f(x, \mathring{y}, \mathring{y}')\, dx.$$

Durch teilweise Integration erhält man für J_2 den Ausdruck

$$J_2 = \frac{1}{2}\int_a^b \left[\mathring{f}_{y'y'}\left(\frac{du}{dx}\right)^2 + 2\mathring{f}_{y'y}\, u\, \frac{du}{dx} + \mathring{f}_{yy}\, u^2\right] dx$$

$$(139) \qquad = \frac{1}{2}\int_a^b \left[p\left(\frac{du}{dx}\right)^2 - ku^2\right] dx - \frac{1}{2}\left(h_1 u^2(a) + h_2 u^2(b)\right),$$

$$k = -\mathring{f}_{yy} + \frac{d}{dx}\mathring{f}_{y'y}, \quad h_1 = f_{y'y}(a, \mathring{y}(a), \mathring{y}'(a)), \quad h_2 = -f_{y'y}(b, \mathring{y}(b), \mathring{y}'(b)),$$

wofür wir, wenn q irgendeine in $[ab]$ analytische und reguläre, wesentlich negative Funktion bezeichnet, schreiben wollen

$$(140) \quad J_2 = \frac{1}{2}\int_a^b \left[p\left(\frac{du}{dx}\right)^2 - qu^2 - (k-q)\, u^2\right] dx\, dy - \frac{1}{2}\left(h_1 u^2(a) + h_2 u^2(b)\right)$$ [2].

Für das Eintreten des Minimums ist notwendig, daß $J_2 \geqq 0$ ausfällt. Daraus folgt in bekannter Weise, wenn insbesondere $u(a) = u(b) = 0$ angenommen wird, daß vor allem die Jacobische Bedingung für das Problem mit festen Grenzen erfüllt sein

---

1) Im Gegensatz zu der in § 2 des ersten Kapitels ebenso bezeichneten Funktion braucht jetzt u für $x = a$ und $x = b$ nicht notwendig zu verschwinden.

2) Die konstanten h_1 und h_2 sind mit den in § 5 des ersten Kapitels zur Abkürzung eingeführten Funktionen h_1 und h_2 nicht zu verwechseln.

muß. Ist λ_1^* der zu dem ersten Randwertproblem gehörige kleinste positive Eigenwert der Differentialgleichung

$$(5)\qquad \frac{d}{dx}\left(p\frac{du}{dx}\right)+\lambda k u = 0,$$

so muß $\lambda_1^* \geqq 1$ sein[1]).

Betrachten wir jetzt diejenigen zu dem Intervalle $[ab]$ gehörigen der Beziehung

$$\int_a^b (k-q)\,\varphi_j^2\,dx + h_1\,\varphi_j^2(a) + h_2\,\varphi_j^2(b) = \frac{\lambda_j}{|\lambda_j|}$$

gemäß normierten Eigenfunktionen der Differentialgleichung

$$(141)\qquad \frac{d}{dx}\left(p\frac{du}{dx}\right)+qu+\lambda(k-q)\,u = 0,$$

die den Grenzbedingungen

$$(142)\quad p(a)\frac{du}{dx}+\lambda h_1 u = 0 \quad \text{für } x=a, \qquad p(b)\frac{du}{dx}-\lambda h_2 u = 0 \quad \text{für } x=b$$

genügen. Es seien λ_j die zugehörigen Eigenwerte. Es gibt unendlich viele Eigenwerte, von denen im allgemeinen unendlich viele positiv und unendlich viele negativ sind. Wie ich neuerdings bewiesen habe[2]), gelten für jede in $[ab]$ stetige Funktion $u(x)$, die abteilungsweise stetige Ableitung hat, die Beziehungen

$$(143)\qquad \begin{aligned}&\int_a^b (k-q)\,u^2\,dx + h_1 u^2(a) + h_2 u^2(b)\\ &= \sum \frac{\lambda_j}{|\lambda_j|}\left[\int_a^b (k-q)\,u\varphi_j\,dx + h_1\,\varphi_j(a)\,u(a) + h_2\,\varphi_j(b)\,u(b)\right]^2,\end{aligned}$$

$$(144)\quad \int_a^b \left[p\left(\frac{du}{dx}\right)^2 - qu^2\right]dx = \sum |\lambda_j| \left[\int_a^b (k-q)\,u\varphi_j\,dx + h_1\,\varphi_j(a)\,u(a) + h_2\,\varphi_j(b)\,u(b)\right]^2.$$

Aus (140), (143) und (144) folgt

1) In dem ersten Kapitel ist dieser Eigenwert mit λ_1 bezeichnet worden.

2) Vgl. L. Lichtenstein, Zur Analysis der unendlich vielen Variablen. Zweite Abhandlung. Reihenentwicklungen nach Eigenfunktionen linearer partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus. Math. Zeitschrift 3 (1919), S. 127—160. Dort wird das analoge Problem der Theorie partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus betrachtet. Im allgemeinen gilt in der Formel (144) das Zeichen $\geqq$. Ist die Funktion $k-q$ in keinem in $[ab]$ gelegenen Intervall identisch gleich Null, so tritt hier das Gleichheitszeichen. Diese Bedingung ist in unserem Falle, wo die Funktionen k und q analytisch und regulär sind, gewiß erfüllt.

$$(145)\quad \boxed{J_2 = \frac{1}{2}\sum |\lambda_j|\left(1-\frac{1}{\lambda_j}\right)\left(\int_a^b (k-q)\,u\varphi_j\,dx + h_1\,\varphi_j(a)\,u(a) + h_2\,\varphi_j(a)\,u(b)\right)^2.}$$

Aus dieser Formel folgt fast unmittelbar:

Damit die zweite Variation $J_2 \geqq 0$ sei, ist notwendig und hinreichend, daß der kleinste positive Eigenwert

$$(146)\qquad \lambda_1 \geqq 1$$

sei.

Die Beziehung (146) ist der im ersten Kapitel für das einfachste Problem bei festem Rand aufgestellten Bedingung ganz analog.

§ 2.

Es sei insbesondere $\lambda_1 > 1$.

Aus (145) folgt

$$(147)\qquad J_2 \geqq \frac{1}{2}\left(1-\frac{1}{\lambda_1}\right)\int_a^b \left[p\left(\frac{du}{dx}\right)^2 - qu^2\right]dx.$$

Sind alle Eigenwerte negativ, in welchem Falle wir $\lambda_1 = +\infty$ setzen, so gilt

$$(148)\qquad J_2 \geqq \frac{1}{2}\int_a^b \left[p\left(\frac{du}{dx}\right)^2 - qu^2\right]dx.$$

Ähnlich wie in § 2 des ersten Kapitels findet man

$$(149)\qquad \Delta J \geqq \varepsilon^2\left(\frac{1}{2}\left(1-\frac{1}{\lambda_1}\right) - \alpha_{33}\,\varepsilon M_1\right)\int_a^b \left[p\left(\frac{du}{dx}\right)^2 - qu^2\right]dx.$$

Für $\lambda_1 > 1$ liegt zunächst ein schwaches Minimum vor.

Ist die Bedingung $\lambda_1 > 1$ bei einer bestimmten Wahl der Funktion $q < 0$ erfüllt, so ist sie natürlich erfüllt, wie man auch $q < 0$ sonst wählen könnte. Man kann sich hiervon auch wie folgt direkt überzeugen. Wäre etwa $\lambda_1 < 1$, wenn man für q eine Funktion $Q < 0$ setzt, so müßte aus Stetigkeitsgründen, wenn man $q + \mu(Q-q)$ für q setzt, für einen Wert des Parameters μ zwischen 0 und 1 der Eigenwert $\lambda_1 = 1$ werden. Die Differentialgleichung

$$(150)\qquad \frac{d}{dx}\left(p\,\frac{dy}{dx}\right) + (q + \mu(Q-q))\,y + (k - q - \mu(Q-q))\,y = 0,$$

oder was dasselbe ist, die Differentialgleichung

$$(151) \qquad \frac{d}{dx}\left(p\,\frac{dy}{dx}\right)+qy+(k-q)\,y = 0$$

hätte eine in $[ab]$ reguläre Lösung, die den Grenzbedingungen

$$p(a)\frac{du}{dx}+h_1 u = 0 \quad \text{für } x=a, \qquad p(b)\frac{du}{dx}-h_2 u = 0 \quad \text{für } x=b$$

genügt. Dies ist indessen nicht möglich, da $\lambda_1 > 1$ der kleinste positive Eigenwert des Problems ist.

Ist $\lambda_1 > 1$, so ist gewiß auch $\lambda_1^* > 1$. Andernfalls würde nämlich J_2 verschwinden können[1]). Da $\lambda_1^* > 1$ ist, so läßt sich $\overset{0}{y}$ in ein Feld von Extremalen einbetten. Wie wir in § 4 des ersten Kapitels bewiesen haben, gibt es eine Zahl $\overline{\overline{\delta}} > 0$, so daß es, wie man die Randwerte y_a und y_b in den Intervallen

$$(152) \qquad |y_a-\overset{0}{y}(a)| \leqq \overline{\overline{\delta}}, \quad |y_b-\overset{0}{y}(b)| \leqq \overline{\overline{\delta}}$$

wählt, eine und nur eine Extremale y gibt, die durch die Punkte (a, y_a), (b, y_b) hindurchgeht. Die Gesamtheit dieser Extremalen bildet eine zweifach unendliche Schar $\mathfrak{S}$. Jede Extremale y der Schar $\mathfrak{S}$ läßt sich in ein durch die Randwerte

$$(153) \qquad y_a+\varepsilon\overline{A}, \quad y_b+\varepsilon\overline{A} \qquad (|\varepsilon| \leqq \varepsilon^{(3)})$$

charakterisiertes Feld einbetten. Das von diesem Felde bestrichene Flächenstück $\overline{\mathfrak{S}}$ enthält für jedes Paar Randwerte y_a, y_b alle Extremalen der Schar $\mathfrak{S}$.

Das Flächenstück, das aus der Gesamtheit aller $\overline{\mathfrak{S}}$ besteht, heiße $\mathfrak{S}'$. Es möge in bekannter Bezeichnungsweise die Weierstraßsche Funktion $E(x, y;\, \dot{p}, \tilde{p})$ für alle $\dot{p}$ in einer Umgebung von $\overset{0}{p} = \frac{d\overset{0}{y}}{dx}$ und alle von $\dot{p}$ verschiedenen $\tilde{p}$ positiv sein. Für hinreichend kleine Werte von $\overline{\overline{\delta}}$ und $\varepsilon^{(3)}$ ist offenbar $E(x, y;\, p, \tilde{p}) > 0$ für alle p in $\mathfrak{S}'$ und alle von p verschiedenen $\tilde{p}$.

Es sei jetzt η irgendeine in $[ab]$ erklärte stetige Funktion, die abteilungsweise stetige Ableitung hat und in dem von der Schar $\mathfrak{S}$ bestrichenen Flächenstück liegt. Es sei η^* diejenige Extremale aus $\mathfrak{S}$, die für $x = a$ und $x = b$ dieselben Werte wie η annimmt. Offenbar liegt η ganz im Innern des vorhin beschriebenen Feldes um η^*. Darum ist

1) Vgl. § 2 des ersten Kapitels.

$$\int_a^b f(x, \eta^*, \eta^{*\prime})\, dx \leqq \int_a^b f(x, \eta, \eta')\, dx. \tag{154}$$

Das Gleichheitszeichen gilt nur für $\eta = \eta^*$. Da, wie wir bereits bewiesen haben, ein schwaches Minimum vorliegt und $\lim\limits_{\bar{\bar{\delta}}=0} \eta^* = \overset{0}{y}$, $\lim\limits_{\bar{\bar{\delta}}=0} \eta^{*\prime} = \overset{0}{y}'$ gilt, so ist ferner für hinreichend kleine $\bar{\bar{\delta}}$

$$\int_a^b f(x, \overset{0}{y}, \overset{0}{y}')\, dx \leqq \int_a^b f(x, \eta^*, \eta^{*\prime})\, dx. \tag{155}$$

Das Gleichheitszeichen gilt nur für $\eta^* = \overset{0}{y}$. Aus (154) und (155) folgt endlich

$$\int_a^b f(x, \overset{0}{y}, \overset{0}{y}')\, dx < \int_a^b f(x, \eta, \eta')\, dx,$$

außer wenn $\eta \equiv \overset{0}{y}$ ist.

Es liegt demnach ein starkes Minimum vor.

§ 3.

Ist $\lambda_1 = 1$, so kann J_2 verschwinden. Dies tritt dann und nur dann ein, wenn $u = \text{const.}\ \varphi_1$ ist[1]. Daraus folgt sofort, daß φ_1 im Innern von $[ab]$ von Null verschieden ist[2]. Demnach muß $\lambda_1^* \geqq 1$ sein. Wäre $\lambda_1^* = 1$, so müßte, da dann alle im Innern von $[ab]$ positiven Lösungen am Rande verschwinden, $\varphi_1(a) = \varphi_1(b) = 0$ sein. Das hätte aber wegen (142) die weiteren Randbedingungen

$$\frac{d\varphi_1(a)}{dx} = 0, \qquad \frac{d\varphi_1(b)}{dx} = 0$$

zur Folge. Nach bekannten Sätzen würde darum $\varphi_1 \equiv 0$ sein. Also ist $\lambda^* > 1$. Die Extremale $\overset{0}{y}$ läßt sich in ein Feld einbetten.

---

1) Die Eigenfunktion φ_1 ist offenbar diejenige in $[ab]$ reguläre Lösung der Differentialgleichung $\frac{d}{dx}\left(p\,\frac{du}{dx}\right) + ku = 0$, die den Grenzbedingungen

$$p(a)\,\frac{du(a)}{dx} + h_1 u(a) = 0, \quad p(b)\,\frac{du(b)}{dx} - h_2 u(b) = 0 \text{ genügt.}$$

2) Es möge im Gegensatz hierzu φ_1 in $[ab]$ verschwinden. Es sei x' ein Punkt, in dem $\varphi_1 \neq 0$ ist und es sei $\xi_1\,\xi_2$ das größte x' enthaltende Intervall, in dessen Innern φ_1 nicht verschwindet. Wir setzen $\varphi_1^* = \varphi_1$ in $[ab] - [\xi_1\,\xi_2]$, $\varphi_1^* = -\varphi_1$ in $[\xi_1\,\xi_2]$. Entgegen der Behauptung des Textes würde für $u = \varphi_1^*$ die zweite Variation $J_2 = 0$ sein.

Ob und unter welchen Bedingungen ein Minimum eintreten kann, bedarf einer näheren Diskussion, die ganz wie in § 5 des ersten Kapitels durchgeführt werden kann[1]). Liegt ein Minimum vor, so ist es stets ein starkes Minimum. Der Beweis läßt sich, da $\overset{0}{y}$ in ein Feld eingebettet werden kann, wie in § 2 erbringen.

Berlin, den 9. Februar 1919.

---

1) Für das Gelingen des Verfahrens war dort die Konvergenz der Reihe $\sum \frac{\varphi_j(x)\,\varphi_j(y)}{|\lambda_j|}$ wesentlich. Auch jetzt ist diese Reihe unbedingt und gleichmäßig konvergent, und es gilt

$$\sum \frac{\varphi_j(x)\,\varphi_j(y)}{|\lambda_j|} = G(x, y),$$

unter $G(x, y)$ die zu den Randbedingungen $\frac{du(a)}{dx} = \frac{du(b)}{dx} = 0$ gehörige Greensche Funktion der Differentialgleichung $\frac{d}{dx}\left(p\,\frac{du}{dx}\right) + qu = 0$ verstanden. Der Beweis läßt sich ganz wie in meiner in der Anmerkung 1 S. 166 unter a) zitierten Abhandlung S. 128 durchführen.

Schraubenpropeller mit geringstem Energieverlust.

Von

Albert Betz.

Mit einem Zusatz von **L. Prandtl.**

Vorgelegt in der Sitzung vom 28. März 1919 durch L. Prandtl.

Wenn man mit einem Tragflügel eines Flugzeuges Auftrieb (senkrecht zur Bewegungsrichtung) erzeugen will, so ist dies nur unter gleichzeitiger Entstehung von Widerstand (entgegengesetzt der Bewegungsrichtung) zu erreichen[1]). Dieser rührt daher, daß zur Erzeugung des Auftriebes der Luft eine nach abwärts gerichtete Geschwindigkeit erteilt werden muß, deren kinetische Energie durch Arbeitsleistung, also durch Überwindung eines Widerstandes aufgebracht werden muß. Die Größe dieses Widerstandes hängt außer von Auftrieb, Spannweite, Fahrgeschwindigkeit und Luftdichte auch von der Verteilung des Auftriebes über die Spannweite ab. Nach einem Satz von Munk[2]) ist sie unter Beibehaltung der übrigen maßgebenden Größen dann ein Minimum, wenn die der Luft erteilte Vertikalgeschwindigkeit an den von der Tragfläche durchlaufenen Stellen konstant ist[3]).

Die Strömung hinter einer Tragfläche mit günstigster Auftriebsverteilung ist demnach so, wie wenn der von der Tragfläche durchlaufene Raum zu einem festen Körper (Brett) erstarrt

1) Vergl. Prandtl, Tragflügeltheorie I. Mitt., Abschn. 10. Göttinger Nachrichten 1918, S. 451.

2) Munk, Isoperimetrische Aufgaben aus der Theorie des Fluges. Dissertation Göttingen 1919.

3) Vergl. Prandtl, Tragflügeltheorie II. Mitt., Abschn. 5. Göttinger Nachrichten 1919, S. 107.

wäre, der sich mit einer bestimmten Geschwindigkeit nach abwärts bewegt.

Bei einem Schraubenpropeller, der dazu dient, ein Drehmoment in Schub umzusetzen, liegen nun sehr verwandte Verhältnisse vor. Zur Erzeugung des Schubes muß Flüssigkeit axial nach hinten beschleunigt werden (Schraubenstrahl) und durch Ausübung des Drehmomentes wird ihr auch eine Drehbewegung erteilt. Es muß deshalb bei der Umsetzung von Drehmoment in Schub außer der Schubarbeit auch noch die kinetische Energie der Flüssigkeit im Schraubenstrahl von der Antriebsmaschine aufgebracht werden. Diese nicht nutzbar zu verwertende Arbeit drückt sich in einem maximalen theoretischen Wirkungsgrad aus, der von den Größen: Schub, Schraubendurchmesser, Fahr- und Drehgeschwindigkeit und Dichte der Flüssigkeit abhängt. Sieht man von der Drehenergie ab, die meist unbedeutend ist, so wird das Maximum des Wirkungsgrades erreicht, wenn die Axialgeschwindigkeit über den ganzen Strahlquerschnitt konstant ist (gewöhnliche Schraubenstrahltheorie). Um dies zu erreichen, muß der Schub über die ganze Schraubenkreisfläche gleichmäßig verteilt sein, was aber wiederum eine gleichmäßige Verteilung der Schraubenflügel oder mit anderen Worten unendlich viele Schraubenflügel voraussetzt.

Bei wirklichen Schrauben ist nur eine endliche Anzahl von Flügeln möglich und deshalb der Schub auf diese Stellen konzentriert. Bei der geringen Flügelbreite kann man sich den Schub eines Schraubenflügels auf eine Linie konzentriert denken (entsprechend der tragenden Linie, welche den Flügel eines Flugzeuges in der Theorie ersetzt[1]) Will man aber die endliche Flügelbreite berücksichtigen, so kann man verhältnismäßig leicht von den unendlich schmalen Flügeln dazu übergehen, indem man durch Integration eine Schar von Linien zu einem Körper zusammenfaßt. Wenn man so den Schub auf einzelne Linien konzentriert, und außerdem die Drehenergie des Strahles in Betracht zieht, so erhält man wesentlich weniger einfache Verhältnisse. Immerhin gibt es auch hierbei eine verhältnismäßig einfache Darstellung für jene hinter einer Schraube erzeugte Bewegung, welche den geringsten Energieverlust, also den größten Wirkungsgrad bei gegebenen äußeren Bedingungen mit sich bringt:

Die Strömung hinter einer Schraube mit geringstem Energieverlust ist so, wie wenn die von jedem Schraubenflügel durchlaufene Bahn (Schrauben-

1) Vergl. Prandtl, Tragflügeltheorie I, 12, a. a. O.

fläche) erstarrt wäre und sich mit einer bestimmten Geschwindigkeit nach hinten verschiebt oder sich mit einer bestimmten Winkelgeschwindigkeit um die Schraubenachse dreht[1]).

Der Satz wurde zunächst als Analogie zu dem entsprechenden oben für die Tragfläche angeführten aufgestellt und hatte eine gewisse Wahrscheinlichkeit für sich. Der Beweis für seine Richtigkeit soll im Folgenden gebracht werden[2]). Der Satz gilt für beliebig geformte Schrauben. Im Interesse einer leichteren Verständlichkeit der Beweisführung soll jedoch zunächst vorausgesetzt werden, daß die Schraubenflügel geradlinig von der Achse nach außen hin verlaufen, also je mit einem Radius zusammenfallen. Nachträglich wird dann der Beweis auch auf beliebige krummlinige Flügel ausgedehnt werden.

Wie bei der Tragfläche sind auch bei der Schraube die auf einen Flügel ausgeübten Kräfte pro Längeneinheit (abgesehen von reinen Widerstandskräften) gleich dem Produkt $\Gamma \cdot c \cdot \varrho$ aus der Zirkulation um das Flügelprofil Γ, der Geschwindigkeit desselben c und der Flüssigkeitsdichte ϱ und sind senkrecht zur Geschwindigkeit c gerichtet[3]). Von der Hinterkante eines Flügels gehen

---

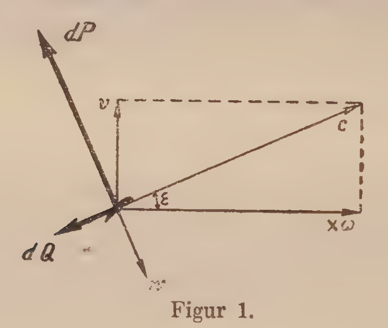
1) Da man eine Schraubenfläche in sich selbst verschieben kann und dies offenbar keinen Einfluß auf die Flüssigkeitsbewegung hat, so kann man diese Bewegung der Schraubenfläche, die sich aus einer Drehung um die Achse und einer Verschiebung in axialer Richtung zusammensetzt, jeder anderen Bewegung der Fläche überlagern. Dadurch kann man eine reine Axialbewegung in eine reine Drehbewegung überführen und umgekehrt, ohne daß dadurch an der Strömung der umgebenden Flüssigkeit etwas geändert wird. Es ist deshalb auch gleichgültig, ob man eine axiale Verschiebung der Schraubenfläche, oder eine Drehung um die Achse betrachtet.

2) Ebenso wie bei den entsprechenden Problemen der Tragflügeltheorie werden auch hier die von den Flügeln verursachten Störungsgeschwindigkeiten als klein gegenüber der Eigengeschwindigkeit der Flügel vorausgesetzt. Der Beweis gilt also streng nur für Schrauben mit kleinem Schub (bezw. kleinem Belastungsgrad $\frac{S}{\varrho F v^2}$, wo $S =$ Schub, $\varrho =$ Luftdichte, $F = \frac{D^2 \pi}{4} =$ Schraubenkreisfläche, $v =$ Fahrgeschwindigkeit). Während bei Tragflügeln diese Voraussetzung stets mit hinreichender praktischer Genauigkeit zutrifft, kommen bei Schraubenpropellern, insbesondere für Schiffe, ziemlich hohe Belastungsgrade vor, so daß hierbei schon merkliche Abweichungen von der hier gegebenen Theorie zu erwarten sind. Weiter bleiben hier ebenso wie bei der Tragflügeltheorie die Reibungsverluste unberücksichtigt, sie sind im übrigen auch praktisch nur von geringer Bedeutung gegenüber den theoretisch erfaßbaren Verlusten.

3) Vergl. Prandtl, Tragflügeltheorie I, 6, a. a. O.

Wirbel aus; pro Längeneinheit so viele, daß ihre Zirkulation $\frac{\partial \Gamma}{\partial x}$ ist (x = Abstand von der Schraubenachse). Sie verlaufen im wesentlichen in der vom Flügel zurückgelegten Bahn (wenn man von ihrer geringen Eigenbewegung absieht [1]), also in Schraubenlinien. In ihrer Gesamtheit bilden sie eine Schraubenfläche. Ihr Geschwindigkeitsfeld ist die durch die Bewegung des Flügels verursachte Strömung.

Hat ein Teil eines Flügels (im Abstand x von der Achse) die Geschwindigkeit $c = \sqrt{v^2 + (x\omega)^2}$ (Fig. 1; v = Fahrgeschwindigkeit, ω = Winkelgeschwindigkeit der Schraube), so würde bei ruhender Flüssigkeit die Relativgeschwindigkeit derselben gegen das Flügelprofil die Größe und Richtung $-c$ haben. Diese erzeugt an dem Flügelelement, dessen Länge dx sein möge, in Verbindung mit der dort herrschenden Zirkulation Γ eine Kraft $dP = \varrho \cdot \Gamma \cdot c \cdot dx$, die senkrecht zur Geschwindigkeit c und zum Radius x steht (Fig. 1).



Figur 1.

Tritt nun infolge des Wirbelfeldes noch eine Störungsgeschwindigkeit w hinzu, die in dem betrachteten Punkte senkrecht zur Bewegungsrichtung c und zum Radius (Flügel), also senkrecht zu der vom Flügel durchlaufenen Schraubenfläche steht, so ergibt dies eine Kraft $dQ = \Gamma \cdot w \cdot dx$, die entgegengesetzt zur Bewegungsrichtung c liegt. Zur Überwindung dieser Kraft dQ muß bei der Geschwindigkeit c eine sekundliche Arbeit $c \cdot dQ$ geleistet werden. Eine Störungsgeschwindigkeit, welche in die Schraubenfläche fällt würde eine Kraft senkrecht zu c mit sich bringen, welche also keine zusätzliche Arbeit verursacht; wir brauchen deshalb bei der Beurteilung der Energieverhältnisse auch nur die Störungsgeschwindigkeit, welche senkrecht zur Schraubenfläche steht, in Betracht zu ziehen.

1) S. Anm. 2 voriger Seite.

Wenn für gegebene Werte von Schub S, Durchmesser $2R$, Geschwindigkeit v, Winkelgeschwindigkeit ω und Flüssigkeitsdichte ϱ, bei einer bestimmten Verteilung der Zirkulation längs der Flügel (längs x) der zusätzliche Energieverbrauch $\int c \cdot dQ$ ein Minimum sein soll, so muß jede kleine Änderung in der Verteilung, welche den gleichen Schub ergibt, in erster Näherung den Energieverbrauch unverändert lassen ($\delta \int c \cdot dQ = 0$ für $S =$ konst.).

Nach Einführen des Wertes für dQ lautet die Minimumsbedingung:

$$\delta \sum_1^n \int_0^R \varrho \cdot \Gamma \cdot w \cdot \frac{v}{\sin \varepsilon} \cdot dx = 0 \text{ [1]}, \qquad (1)$$

wobei noch $c = \frac{v}{\sin \varepsilon}$ gesetzt ist, mit $\operatorname{tg} \varepsilon = \frac{v}{x\omega}$ ($\varepsilon =$ Steigungswinkel der Schraube).

Von der Kraft dP, welche durch die ungestörte Geschwindigkeit c am Flügelelement mit der Zirkulation Γ erzeugt wird, dient die in die Achsrichtung fallende Komponente (Fig. 1)

$$dP \cdot \frac{x\omega}{c} = \varrho \cdot \Gamma \cdot x \cdot \omega \cdot dx$$

zur Schuberzeugung. Die Anteile, welche von den Störungsgeschwindigkeiten bezw. den damit verbundenen Zusatzkräften herrühren, kommen dagegen als kleine Größen zweiter Ordnung nicht in Betracht (vergl. Anm. 2 S. 195). Die Nebenbedingung, daß der Schub konstant sein soll, nimmt demnach die Form an

$$S = \sum_1^n \int_0^R \varrho \cdot \Gamma \cdot x \cdot \omega \cdot dx = \text{konst.} \text{ [1]}. \qquad (2)$$

Vermehren wir an der Stelle x_i (zwischen x_i und $x_i + dx$) den Wert Γ um den Betrag $\frac{\varDelta}{x_i}$ und vermindern ihn an der Stelle x_k (zwischen x_k und $x_k + dx$; die Strecke dx soll in beiden Fällen dieselbe sein) um den Betrag $\frac{\varDelta}{x_k}$, so wird dadurch der Schub nicht geändert, da er an der einen Stelle um $\varDelta \cdot \varrho \cdot \omega \cdot dx$ vermehrt, an der anderen aber um den gleichen Betrag vermindert wird.

---

1) Das $\sum_1^n$-Zeichen soll andeuten, daß die Integration bei sämtlichen Flügeln (es mögen n Flügel sein) auszuführen und die Summe dieser n Integrale zu bilden ist.

Soll auch der Wert des Energieintegrales $\sum_1^n \int_0^R \varrho \cdot \Gamma \cdot w \cdot \frac{v}{\sin \varepsilon} dx$ unverändert bleiben, so müssen sich die durch die Änderung von Γ erzielten Zusatzgrößen der Energie gegenseitig aufheben.

Wir müssen nun zunächst untersuchen, wie groß die Änderung der Verlustenergie ist, wenn wir an irgend einer Stelle längs einer Strecke dx eine Zirkulation $d\Gamma$ hinzufügen. Diese Änderung zerfällt in zwei Teile. Einmal wird die Verlustenergie dadurch verändert, daß an der Stelle, wo $d\Gamma$ hinzugefügt wurde, bereits eine Störungsgeschwindigkeit herrscht, welche zusammen mit dem hinzugefügten Zirkulationselement eine Widerstand bildende Kraft erzeugt; andererseits wird aber durch das neue Zirkulationselement und die von seinen Enden ausgehenden beiden Schraubenwirbel auch die Störungsgeschwindigkeit an allen übrigen Stellen geändert und damit auch die Verlustarbeit, welche an den übrigen mit Zirkulation versehenen Flügelteilen auftritt.

Wir können die Energieverhältnisse aber auch noch von einem anderen Standpunkte aus betrachten. Die gesamte beim Arbeiten der Schraube verlorene Energie ist hinter der Schraube in Form von kinetischer Energie vorhanden, welche in dem von den Schraubenwirbeln herrührenden Strömungsfeld steckt[1]). Wir können deshalb die sekundliche Verlustarbeit anstatt an der Stelle, wo sie entsteht (an den Zirkulation führenden Flügelteilen) auch weit hinter der Schraube dadurch bestimmen, daß wir die kinetische Energie zwischen zwei im Abstand v voneinander befindlichen Ebenen senkrecht zur Schraubenachse ermitteln. Da nämlich der Schraubenstrahl infolge der Fortbewegung der Schraube in jeder Sekunde um den Betrag v länger wird, wird die Energie desselben um den zwischen den beiden erwähnten Ebenen befindlichen Betrag größer. Befinden sich diese Ebenen in großer Entfernung von der Schraube, so hängt die Geschwindigkeitsverteilung nur von der Stärke und Anordnung der von den Flügelhinterkanten ausgehenden schraubenförmigen Wirbel ab, da deren Anfang (die Schraubenflügel) wegen der großen Entfernung keinen Einfluß mehr hat. Es ist daher auch gleichgültig, an welcher Stelle der einzelnen Schraubenwirbel deren Anfang liegt, wenn er nur weit genug von den Kontrollebenen weg ist. Die Verlustenergie bleibt daher auch unverändert, wenn wir das kleine hinzugefügte Zirkulationselement $d\Gamma \cdot dx$ längs der von seinen End-

1) Von Reibungsverlusten ist wie bereits erwähnt abgesehen. Vergl. Anmerkung 2, S. 195.

punkten ausgehenden Schraubenwirbel, d. i. längs seiner Bahn, verschieben. Bei einer Verschiebung nach hinten, wird aber die Verteilung der Verlustenergie auf die Entstehungsstellen geändert. Die von den nicht verschobenen Teilen der Flügel herrührende Störungsgeschwindigkeit ist hinter dem Propeller größer, da sich die Schraubenwirbel von dort aus nach vorn und nach hinten erstrecken, während sie in der Flügelebene selbst nur nach einer Seite hin vorhanden sind; umgekehrt nimmt vor dem Propeller der Störungseinfluß rasch mit der Entfernung ab. Es wird daher bei der Verschiebung nach hinten der Anteil größer, welcher durch Zusammenwirken der von den Flügeln herrührenden Störungsgeschwindigkeit mit dem susätzlichen Zirkulationselement entsteht, während die Störung, welche dieses Zirkulationselement mit seinen Wirbeln an den Flügeln verursacht, mit zunehmender Entfernung immer kleiner wird und damit auch die dadurch verursachte Verlustarbeit. Wenn wir das zusätzliche Element genügend weit nach hinten verschieben, so verschwindet der letztere Anteil vollständig und die ganze durch das Hinzufügen des Elementes verursachte Verluständerung ist durch das Zusammenwirken der vom ganzen Propeller herrührenden Störungsgeschwindigkeit mit dem Zirkulationselement gegeben. Bezeichnen wir mit w' die Komponente der Störungsgeschwindigkeit, welche an der Stelle des verschobenen Zirkulationselementes (also weit hinter der Schraube) senkrecht zu diesem und zu seiner Bahn, also senkrecht zu der von ihm beschriebenen Schraubenfläche steht, so ist, wie wir weiter oben sahen, die dadurch entstehende sekundliche Verlustarbeit $\varrho \cdot d\Gamma \cdot w' \cdot c \cdot dx$ und dies ist, wie wir eben zeigten, die ganze durch das Hinzufügen von $d\Gamma$ verursachte Mehrarbeit.

Nun können wir auf unsere Minimumbetrachtung zurückkommen. Darnach durfte sich die Verlustenergie nicht ändern, wenn wir an der einen Stelle der Flügel im Abstand x_i von der Achse längs einer Strecke dx Zirkulation von der Stärke $\frac{\Delta}{x_i}$ hinzufügten und an einer anderen Stelle im Abstand x_k von der Achse längs einer ebenso großen Strecke Zirkulation von der Stärke $\frac{\Delta}{x_k}$ wegnahmen. Die durch das Hinzufügen und Wegnehmen dieser Zirkulation verursachte Energieänderung läßt sich nach dem oben Auseinandergesetzten durch folgenden Ausdruck angeben, wobei wieder $c = \frac{v}{\sin \varepsilon}$ gesetzt ist (ε = Steigungswinkel

der Schraubenfläche in dem betreffenden Radius). Die Indices i und k geben an, daß die betreffenden Werte an den Stellen x_i und x_k gemeint sind.

$$\varrho \cdot \frac{\Delta}{x_i} \cdot w_i' \cdot \frac{v}{\sin \varepsilon_i} dx - \varrho \cdot \frac{\Delta}{x_k} \cdot w_k' \cdot \frac{v}{\sin \varepsilon_k} dx = 0,$$

daraus ergibt sich

$$\frac{w_i'}{x_i \omega} \cdot \frac{1}{\sin \varepsilon_i} = \frac{w_k'}{x_k \omega} \cdot \frac{1}{\sin \varepsilon_k} = \text{Konst.} = C\,^{1)}$$

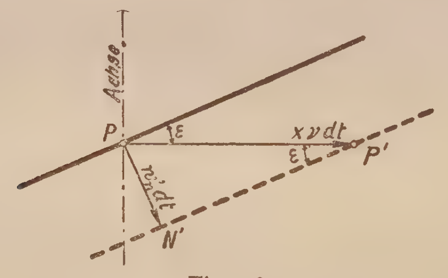
oder

$$\frac{w'}{x\omega} = C \cdot \sin \varepsilon. \tag{3}$$

Diese Bedingung wird aber durch Drehen einer starren Schraubenfläche vom Radius R und der Steigung $H = 2\pi \frac{v}{\omega}$ (Bahn des Flügels) gleichfalls erzeugt. Durch Drehung einer solchen Fläche mit der Winkelgeschwindigkeit ν wird nämlich ein im Abstand x von der Achse befindliches Flüssigkeitsteilchen, abgesehen von seitlichen Verschiebungen, normal zur Schraubenfläche mit der Geschwindigkeit

$$w_n' = x \cdot \nu \cdot \sin \varepsilon \tag{4}$$

verschoben, wie man sofort aus Fig. 2 ablesen kann, die einen Schnitt durch die Schraubenfläche senkrecht zum Radius in zwei



Figur 2.

um die Zeit dt auseinanderliegenden Zeitpunkten darstellt. Dies ist aber, wenn $\nu = C \cdot \omega$ gewählt wird, gerade der Wert von w' der Minimumsbedingung. Da außerdem die Richtung von w_n' mit der von w' übereinstimmt und im übrigen beide Bewegungen Potentialbewegungen sind, so ist damit gezeigt, daß die beiden Vor-

1) Die Winkelgeschwindigkeit ω ist eingeführt, um die Konstante C dimensionslos zu machen.

gänge, hervorgerufen einerseits durch den Schraubenpropeller von günstigstem Wirkungsgrad, andererseits durch die sich drehende Schraubenfläche identisch sind.

Wir hatten bisher der größeren Einfachheit halber vorausgesetzt, daß die Propellerflügel mit Radien zusammenfallen. Der Beweis läßt sich aber mit ganz geringen Änderungen auch für beliebig geformte Flügel führen, nur werden dabei die Verhältnisse etwas weniger übersichtlich. Wir betrachten wieder an irgend einer Stelle der Flügel im Abstand x von der Achse längs einer Strecke dl ein Zirkulationselement Γ, wobei nun dl mit dem Radius einen Winkel bilden möge. Wir können diese Strecke dl in die drei Komponenten zerlegen, dx in Richtung des Radius, dy in Richtung der Bahngeschwindigkeit c und dz in Richtung senkrecht dazu. Das Zirkulationselement von der Länge dl läßt sich dann durch drei Elemente von der gleichen Stärke längs der drei Strecken dx, dy und dz ersetzen. Für den Schub liefert nur das Element dx einen Beitrag, da die anderen beiden Zirkulationskomponenten mit der Geschwindigkeit c zusammen nur Kräfte in radialer Richtung ergeben. Dieser Beitrag ergibt sich demnach genau wie bei geradlinigen Flächen zu:

$$dS = \varrho \cdot \Gamma \cdot x \cdot \omega \cdot dx.$$

Zur Verlustarbeit tragen nur die beiden senkrecht zu c liegenden Elemente dx und dz bei, da das in die Bewegungsrichtung c fallende Element dy nur eine Kraft senkrecht dazu ergibt, welche mit c zusammen keine Arbeit liefert. Wir fassen diese beiden Elemente dx und dz zweckmäßig zu einem besonderen Element ds zusammen, welches also die Projektion von dl auf die zur Bewegungsrichtung c senkrechte Ebene ist. Mit diesem Element zusammen liefert eine Störungsgeschwindigkeit w, die senkrecht zur Bewegungsrichtung und zum Element ds, also senkrecht zu der bei der Bewegung beschriebenen Schraubenfläche[1]) steht, eine Verlustarbeit $\varrho \cdot \Gamma \cdot w \cdot c \cdot ds$. Störungsgeschwindigkeiten, die in diese Schraubenfläche fallen, ergeben nur Kräfte senkrecht zu c, also keine Verlustarbeit. Der Winkel, den ds mit dem Radius, also mit dx bildet, ist derselbe wie der, den die von ds bezw. dl beschriebene Schraubenfläche[1]) mit dem Radius bildet. Wir wollen ihn mit j bezeichnen. Dann ist $ds = \dfrac{dx}{\cos j}$ und unsere Mini-

1) Die Strecken dl und ds beschreiben bei der Bewegung dieselbe Schraubenfläche.

mumsbedingung nimmt jetzt die Form an (wobei wieder $c = \frac{v}{\sin \varepsilon}$ eingeführt ist):

$$\delta \sum_1^n \int_0^R \varrho \cdot \Gamma \cdot w \cdot \frac{v}{\sin \varepsilon} \cdot \frac{1}{\cos j} \cdot dx = 0 \tag{5}$$

mit der Nebenbedingung

$$S = \sum_1^n \int_0^R \varrho \cdot \Gamma \cdot x \cdot \omega \cdot dx = \text{konst.} \tag{6}$$

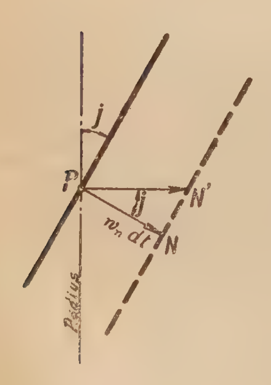
Die Nebenbedingung hat sich gegenüber dem früheren einfacheren Problem überhaupt nicht geändert und in der Variationsgleichung ist nur der Faktor $\frac{1}{\cos j}$ hinzugekommen, der bei radial stehenden Flügeln 1 ist.

Der weitere Gedankengang ist genau der gleiche wie vorher. Wir fügen an einer Stelle x wieder eine zusätzliche Zirkulation im Betrag $\frac{\varDelta}{x}$ hinzu, verschieben sie so weit nach hinten, bis die ganze durch das Element hervorgerufene Verlustenergie an dem Element selbst entsteht und gegeben ist durch den Ausdruck

$$\varrho \cdot \frac{\varDelta}{x} \cdot w' \cdot c \cdot ds = \varrho \cdot \frac{\varDelta}{x} \cdot w' \cdot \frac{v}{\sin \varepsilon} \cdot \frac{1}{\cos j} \cdot dx,$$

wobei w' wieder die weit hinter dem Propeller senkrecht zur Schraubenfläche herrschende Störungsgeschwindigkeit auf der durch den Punkt x gehenden Schraubenlinie ist. In gleicher Weise wie vorher finden wir nun für w' die Bedingung:

$$\frac{w'}{x\omega} = C \cdot \sin \varepsilon \cdot \cos j. \tag{7}$$



Figur 3.

Dies ist aber auch wieder gerade die Geschwindigkeit, welche man durch Drehen der von den Schraubenflügeln beschriebenen Schraubenflächen erhält, wie man aus Fig. 3 ersehen kann. Dort sind wieder zwei aufeinander folgende Lagen der Schraubenfläche gezeichnet und zwar in einem Schnitt senkrecht zur Bewegungsrichtung c, der den Radius enthält. In ihm ist die Normale PN zur Schraubenfläche und der Winkel j in wahrer Größe vorhanden. Daraus ist ohne weiteres zu ersehen, daß $PN = PN' \cos j$

ist (PN' soll senkrecht zum Radius sein) und aus der Figur 2, welche einen Schnitt senkrecht zum Radius darstellt, ist, wie bereits früher geschehen, abzulesen, daß $PN' = x \cdot \nu \cdot \sin\varepsilon \cdot dt$ ist, wobei wieder ν die Winkelgeschwindigkeit des Systems, $x\nu\, dt = PP'$ also die Verschiebung des Punktes P in der Zeit dt ausdrückt. Die Normalgeschwindigkeit $w_n = \frac{PN}{dt}$ ergibt sich demnach zu:

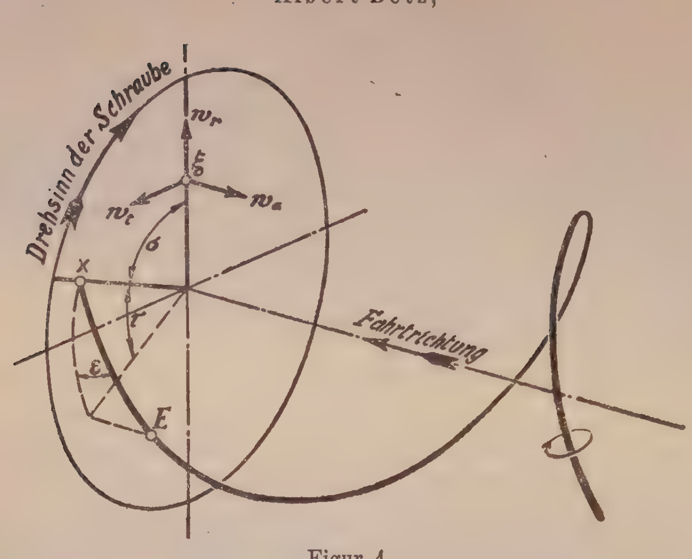
$$w_n = x\nu \sin\varepsilon \cos j, \tag{8}$$

was wieder mit der Geschwindigkeit der Minimumsaufgabe nach Größe und Richtung übereinstimmt, wenn $\nu = C\omega$ gesetzt wird.

Damit ist der am Anfang dieser Untersuchung aufgestellte Satz über die mit geringstem Verlust verknüpfte Strömung hinter einer Schraube auch für beliebig geformte Flügel bewiesen. Die gegebene Beweisführung ist im wesentlichen auf physikalischer Anschauung aufgebaut; es dürfte vielleicht von Interesse sein, zu zeigen, wie man den Beweis auch auf mehr formalem Wege führen kann, zumal da hierbei durch die aufgestellten mathematischen Beziehungen teilweise ein tiefer gehender Einblick in die Vorgänge in der Schraube gewährt wird. Um jedoch nicht zu weitläufig zu werden, sollen nur Schrauben mit geradlinigen radial stehenden Flügeln, die außerdem in einer Ebene liegen, in Betracht gezogen werden. Die Beweisführung für allgemeinere Formen bietet hierbei ebenfalls keine Schwierigkeit, sie würde aber merklich umständlicher ausfallen.

Betrachten wir zunächst die Geschwindigkeit, welche ein von einem Punkte x ausgehender Schraubenwirbel von der Stärke $\Gamma(x)$ an einer Stelle ξ erzeugt. Der Steigungswinkel der Schraubenlinie sei mit $\varepsilon(x)$ bezeichnet $\left(\operatorname{tg}\varepsilon(x) = \frac{v}{x\omega}\right)$. Der Winkel, den die beiden Radien bilden, auf denen die Punkte x und ξ liegen, sei σ (Fig. 4). Man kann diese Geschwindigkeit dadurch ermitteln, daß man den Einfluß eines Elementes E (Fig. 4) des Schraubenwirbels auf den Punkt ξ mittelst des Biot-Savart'schen Gesetzes [1]) angibt und daraus durch Integration über die Länge des Wirbels (von $\tau = 0$ bis $\tau = \infty$; τ ist der Winkel, den die beiden durch E und x gehenden Ebenen einschließen, welche sich längs der Schraubenachse schneiden) den Einfluß des ganzen Wirbels findet.

1) Die Gesetze des Geschwindigkeitsfeldes eines Wirbels sind dieselben wie die des magnetischen Feldes eines elektrischen Stromes.



Figur 4.

Dabei ergibt sich jedoch die Schwierigkeit, daß dieses Integral sich nicht in geschlossener Form auswerten läßt; es wird deshalb im folgenden immer in der Form eines unausgewerteten Integralausdruckes verwandt werden.

Zerlegt man die Geschwindigkeit, die der vom Punkte x ausgehende Schraubenwirbel an der Stelle ξ erzeugt, in die drei Komponenten w_a in axialer, w_t in tangentialer und w_r in radialer Richtung, so sind diese drei Geschwindigkeitskomponenten gegeben durch die Ausdrücke:

$$w_a = \frac{\Gamma(x)}{4\pi} \cdot x \cdot \int_0^\infty N^{-\frac{3}{2}} [x - \xi \cos(\tau + \sigma)] \, d\tau \tag{9}$$

$$w_t = \frac{\Gamma(x)}{4\pi} \cdot \frac{v}{\omega} \cdot \int_0^\infty N^{-\frac{3}{2}} [\xi - x \cos(\tau + \sigma) - x\tau \sin(\tau + \sigma)] \, d\tau \tag{10}$$

$$w_r = \frac{\Gamma(x)}{4\pi} \cdot \frac{v}{\omega} \cdot \int_0^\infty N^{-\frac{3}{2}} [x \sin(\tau + \sigma) - x \cdot \tau \cdot \cos(\tau + \sigma)] \, d\tau \tag{11}$$

wobei zur Abkürzung gesetzt ist

$$x^2 + \xi^2 - 2x\xi \cos(\tau + \sigma) + \frac{v^2}{\omega^2} \tau^2 = N \tag{12}$$

($\sqrt{N}$ ist der Abstand des Elementes E vom Punkte ξ).

Für unsere Betrachtungen brauchen wir nur die Komponente, welche im Punkte ξ senkrecht zur Schraubenfläche steht, die vom

Flügel, dem der Punkt ξ angehört, durchlaufen wird. Hierzu liefert w_r gar keinen Beitrag; w_a und w_t liefern die Beiträge

$$w_a \cdot \frac{\xi\omega}{c(\xi)} \quad \text{und} \quad -w_t \cdot \frac{v}{c(\xi)}$$

(dabei ist $c(\xi) = \sqrt{v^2 + (\xi\omega)^2}$, vergl. Fig. 1). Herrscht an der Stelle ξ die Zirkulation $\Gamma(\xi)$ so verursacht der Wirbel die schädliche Arbeit pro Längeneinheit

$$\varrho \cdot \Gamma(\xi) \left\{ w_a \cdot \frac{\xi\omega}{c(\xi)} - w_t \cdot \frac{v}{c(\xi)} \right\} \cdot c(\xi)$$

$$= \varrho \cdot \frac{\Gamma(x) \cdot \Gamma(\xi)}{4\pi} \cdot \int_0^\infty N^{-\frac{3}{2}} \left\{ [x - \xi \cos(\tau + \sigma)] \cdot x \cdot \xi \cdot w \right.$$

$$\left. + [x \cos(\tau + \sigma) - \xi + x\tau \sin(\tau + \sigma)] \frac{v^2}{\omega} \right\} d\tau$$

$$= \varrho \cdot \frac{\Gamma(x) \cdot \Gamma(\xi)}{4\pi} \cdot \Psi(x, \xi, \sigma). \tag{13}$$

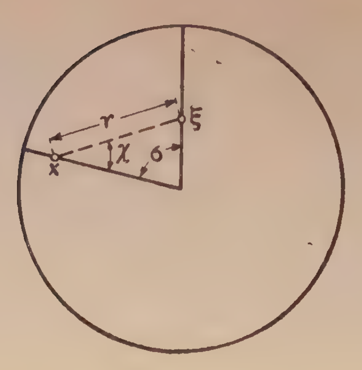
Denken wir uns nun einen Propeller von der Art, daß nur an der Stelle zwischen x und $x + dx$ die Zirkulation $\Gamma(x)$ und an der Stelle zwischen ξ und $\xi + d\xi$ die Zirkulation $\Gamma(\xi)$ herrscht, so gehen einerseits an den Stellen x und $x + dx$ Schraubenwirbel von entgegengesetztem Drehsinn mit der Zirkulation $-\Gamma(x)$ bezw. $+\Gamma(x)$ aus; andererseits gehen entsprechende Wirbel mit der Zirkulation $-\Gamma(\xi)$ und $+\Gamma(\xi)$ von den Stellen ξ und $\xi + d\xi$ aus. Der schädliche Einfluß, den die beiden bei x und $x + dx$ abgehenden Wirbel an der Stelle ξ verursachen, ist nach obiger Formel (13) gegeben durch

$$\varrho \cdot \frac{\Gamma(x) \cdot \Gamma(\xi)}{4\pi} \cdot \frac{\partial \Psi(x, \xi, \sigma)}{\partial x} \cdot dx \cdot d\xi. \tag{14}$$

(Die Differentiation nach x rührt daher, daß es sich um die Differenzwirkung zweier im Abstand dx befindlicher Wirbel handelt). Entsprechend ist die Wirkung der bei ξ und $\xi + d\xi$ abgehenden Wirbel an der Stelle x

$$\varrho \cdot \frac{\Gamma(x) \cdot \Gamma(\xi)}{4\pi} \cdot \frac{\partial \Psi(\xi, x, -\sigma)}{\partial \xi} \cdot dx \cdot d\xi. \tag{15}$$

Mit $\Psi(\xi, x, -\sigma)$ soll dabei jene Funktion bezeichnet werden, in welche $\Psi(x, \xi, \sigma)$ durch Vertauschen von x und ξ einerseits und von σ und $-\sigma$ anderseits übergeht. Außerdem ruft auch das Wirbelstückchen $\Gamma(x)\,dx$ an der Stelle ξ eine axial gerichtete Geschwindigkeit hervor, deren Größe man leicht mittels des



Figur 5.

Biot-Savart'schen Gesetzes aus Fig. 5 ableiten kann. Sie beträgt:

$$-\frac{\Gamma(x)}{r^2}\sin\chi\, dx = -\frac{\Gamma(x)}{r^3}\xi\sin\sigma\, dx, \tag{16}$$

wobei $r = \sqrt{x^2+\xi^2-2x\xi\cos\sigma}$ der Abstand der Punkte x und ξ ist. Die dadurch verursachte Verlustarbeit beträgt (vergl. oben die durch w_a verursachte Arbeit)

$$-\varrho\cdot\frac{\Gamma(x)\Gamma(\xi)}{4\pi}\cdot\frac{\xi\sin\sigma}{r^3}\cdot\frac{\xi\omega}{c(\xi)}\cdot c(\xi)\cdot dx\cdot d\xi$$

$$= -\varrho\cdot\frac{\Gamma(x)\Gamma(\xi)}{4\pi}\cdot\frac{\xi^2\omega\sin\sigma}{r^3}\cdot dx\cdot d\xi. \tag{17}$$

Entsprechend erzeugt das Wirbelstückchen $\Gamma(\xi)d\xi$ an der Stelle x die Verlustarbeit

$$-\varrho\,\frac{\Gamma(x)\Gamma(\xi)}{4\pi}\cdot\frac{x^2\omega\sin(-\sigma)}{r^3}\cdot dx\cdot d\xi. \tag{18}$$

Der gesamte durch das Zusammenwirken der beiden Wirbelstückchen $\Gamma(x)dx$ und $\Gamma(\xi)d\xi$ nebst den zugehörigen Schraubenwirbeln verursachte Verlust beträgt

$$d^2W = \varrho\cdot\frac{\Gamma(x)\Gamma(\xi)}{4\pi}\left[\frac{\partial\Psi(x,\xi,\sigma)}{\partial x} - \frac{\xi^2\omega}{r^3}\sin\sigma + \frac{\partial\Psi(\xi,x,-\sigma)}{\partial\xi} - \frac{x^2\omega}{r^3}\sin(-\sigma)\right]\cdot dx\cdot d\xi$$

$$= \varrho\cdot\frac{\Gamma(x)\Gamma(\xi)}{4\pi}\left[\Xi(x,\xi,\sigma)+\Xi(\xi,x,-\sigma)\right]\cdot dx\cdot d\xi. \tag{19}$$

Dabei ist

$$\Xi(x,\xi,\sigma) = \frac{\partial\Psi(x,\xi,\sigma)}{\partial x} - \frac{\xi^2\omega}{r^3}\sin\sigma$$

der Anteil des Klammerausdruckes, der mit dem Faktor

$$\varrho\cdot\frac{\Gamma(x)\,\Gamma(\xi)}{4\pi}\cdot dx\cdot d\xi$$

zusammen den Teil der Verlustarbeit angibt, welcher an der Stelle ξ durch die Einwirkung der dort herrschenden Störungsgeschwindigkeit auf die Zirkulation $\Gamma(\xi)$ entsteht. Entsprechend stellt

$$\Xi(\xi,x,-\sigma)=\frac{\partial\Psi(\xi,x,-\sigma)}{\partial\xi}-\frac{x^2\omega}{r^3}\sin(-\sigma)$$

den für die Stelle x in Frage kommenden Ausdruck dar.

Die beiden Ausdrücke $\Xi(x,\xi,\sigma)$ und $\Xi(\xi,x,-\sigma)$ gehen ineinander über, wenn man einerseits x und ξ und andererseits σ und $-\sigma$ vertauscht. Wir müssen nun diese Funktion Ξ genauer untersuchen, wobei sich gewisse Symmetrieeigenschaften ergeben, welche für unseren Beweis wesentlich sind.

Nach Gleichung (13) ist

$$\Psi(x,\xi,\sigma)=\int_0^\infty N^{-\frac{3}{2}}\Big\{[x-\xi\cos(\tau+\varrho)]\,x\,\xi\,\omega$$
$$+[x\cos(\tau+\sigma)-\xi+x\tau\cos(\tau+\sigma)]\,\frac{v^2}{\omega}\Big\}\,d\tau, \tag{20}$$

wobei N wieder die in Gleichung (12) angegebene Bedeutung hat. Da sich hieraus $\frac{\partial N}{\partial x}=2[x-\xi\cos(\tau+\sigma)]$ ergibt, wird

$$\frac{\partial\Psi(x,\xi,\sigma)}{\partial x}$$
$$=\int_0^\infty N^{-\frac{3}{2}}\Big\{2x\,\xi\omega-\xi^2\omega\cos(\tau+\sigma)+\frac{v^2}{\omega}\,[\cos(\tau+\sigma)+\tau\sin(\tau+\sigma)]\Big\}\,d\tau$$
$$-3\int_0^\infty N^{-\frac{5}{2}}\Big\{[x-\xi\cos(\tau+\sigma)]\,x\xi\omega+[x\cos(\tau+\sigma)-\xi+x\tau\sin(\tau+\sigma)]\,\frac{v^2}{\omega}\Big\}[x-\xi\cos(\tau+\sigma)]\,d\tau$$
$$=\int_0^\infty N^{-\frac{3}{2}}\Big\{2x\,\xi\,\omega+\frac{v^2}{\omega}\,[\cos(\tau+\sigma)+\tau\sin(\tau+\sigma)]\Big\}\,d\tau$$
$$-\int_0^\infty N^{-\frac{3}{2}}\,\xi^2\omega\cos(\tau+\sigma)\,d\tau$$
$$-3\int_0^\infty N^{-\frac{5}{2}}\Big\{[x-\xi\cos(\tau+\sigma)]\,[x\cos(\tau+\sigma)-\xi]\,\frac{v^2}{\omega}-2\,(x\,\xi)^2\,\omega\cos(\tau+\sigma)\Big\}\,d\tau$$
$$-3\int_0^\infty N^{-\frac{5}{2}}\Big\{[x^2+\xi^2\cos^2(\tau+\sigma)]\,x\,\xi\,\omega+[x-\xi\cos(\tau+\sigma)]\,x\,\tau\,\frac{v^2}{\omega}\sin(\tau+\sigma)\Big\}\,d\tau. \tag{21}$$

Nun kann man das zweite dieser vier Integrale

$$\int_0^\infty \frac{\xi^2 \omega \cos(\tau+\sigma)}{N^{\frac{3}{2}}} d\tau$$

partiell integrieren mit $\cos(\tau+\sigma)$ als integrablen Faktor, wobei zu beachten ist, daß

$$\frac{\partial N}{\partial \tau} = 2\left[x\xi \sin(\tau+\sigma) + \frac{v^2}{\omega^2}\tau\right]$$

und daß für die Grenze $\tau = 0$ $N = r^2$ wird.

$$-\int_0^\infty \frac{\xi^2 \omega \cos(\tau+\sigma)}{N^{\frac{3}{2}}} d\tau$$

$$= \frac{\xi^2 \omega}{r^3} \sin\sigma - 3\int_0^\infty \frac{\xi^2 \omega \sin(\tau+\sigma)}{N^{\frac{5}{2}}}\left[x\xi\sin(\tau+\sigma) + \frac{v^2}{\omega^2}\tau\right] d\tau \qquad (22)$$

Der Integralausdruck rechts in der letzten Gleichung ergibt mit dem letztem Gliede der Gleichung (21) zusammengenommen

$$-3\int_0^\infty N^{-\frac{5}{2}}\Big\{[x^2 + \xi^2\cos^2(\tau+\sigma) + \xi^2\sin^2(\tau+\sigma)]\,x\xi\omega$$
$$+ [x^2 - x\xi\cos(\tau+\sigma) + \xi^2]\frac{v^2}{\omega}\tau\sin(\tau+\sigma)\Big\} d\tau$$
$$= -3\int_0^\infty N^{-\frac{5}{2}}\Big\{[x^2+\xi^2][x\xi\omega + \frac{v^2}{\omega}\tau\sin(\tau+\sigma)]$$
$$- x\xi\frac{v^2}{\omega}\tau\sin(\tau+\sigma)\cos(\tau+\sigma)\Big\} d\tau.$$

Schreiben wir nun die vollständige Funktion $\Xi(x, \xi, \sigma)$ an, so kommt zu den eben umgeformten Gliedern des Ausdruckes für $\frac{\partial \Psi(x,\xi,\sigma)}{\partial x}$ noch das Glied $-\frac{\xi^2\omega}{r^3}\sin\sigma$ aus Gl. (16) hinzu. Dieses hebt sich aber mit dem bei der partiellen Integration Gleichung 22) entstandenen Gliede $+\frac{\xi^2\omega}{r^3}\sin\sigma$ gerade weg. Wir erhalten deshalb für Ξ den Ausdruck:

$$\Xi(x, \xi, \sigma)$$
$$= \int_0^\infty N^{-\frac{3}{2}}\Big\{2x\xi\omega + \frac{v^2}{\omega}[\cos(\tau+\sigma) + \tau\sin(\tau+\sigma)]\Big\} d\tau$$
$$-3\int_0^\infty N^{-\frac{5}{2}}\Big\{[x - \xi\cos(\tau+\sigma)][x\cos(\tau+\sigma) - \xi]\frac{v^2}{\omega} - 2(x\xi)^2\omega\cos(\tau+\sigma)$$
$$+[x^2+\xi^2][x\xi\omega + \frac{v^2}{\omega}\tau\sin(\tau+\sigma)]$$
$$- x\xi\frac{v^2}{\omega}\tau\sin(\tau+\sigma)\cos(\tau+\sigma)\Big\} d\tau. \qquad (23)$$

Aus dieser Darstellung der Funktion Ξ ist zunächst ersichtlich, daß sie in x und ξ symmetrisch ist, daß sie also ihren Wert nicht ändert, wenn man x und ξ miteinander vertauscht. Die Funktion $\Xi(\xi, x, -\sigma)$ unterscheidet sich demnach von ihr nur noch durch das Vorzeichen von σ und wenn $\sigma = 0$ ist, d. h. wenn x und ξ auf demselben Radius liegen, so ist $\Xi(x, \xi, 0) = \Xi(\xi, x, 0)$, was wir in folgendem Satz ausdrücken können:

Liegen zwei Elemente eines Schraubenflügels auf einem und demselben Radius in der Entfernung x und ξ von der Schraubenachse, so ist die Verlustarbeit, welche an der Stelle ξ durch die von der Zirkulation an der Stelle x herrührende Störungsgeschwindigkeit verursacht wird, gleich der Verlustarbeit, welche an der Stelle x durch die von der Zirkulation an der Stelle ξ herrührende Geschwindigkeit erzeugt wird.

Weiter ist aus Gleichung 23 zu erkennen, daß die Funktion $\Xi(x, \xi, \sigma)$ und entsprechend auch $\Xi(\xi, x, -\sigma)$ ihren Wert nicht ändert, wenn man gleichzeitig die Vorzeichen von τ und σ vertauscht, da immer nur die Verbindungen $\cos(\tau+\sigma)$, $\tau \cdot \sin(\tau+\sigma)$ und τ^2 vorkommen. Es ist demnach, wenn wir noch die eben erwähnte Vertauschbarkeit von x und ξ mit benützen und die Ausdrücke unter dem Integralzeichen mit Ξ' bezeichnen:

$$\Xi(x, \xi, \sigma) = \int_0^\infty \Xi'(x, \xi, \tau+\sigma)\,d\tau = \int_0^\infty \Xi'(\xi, x, -\tau-\sigma)\,d\tau.$$

Da aber

$$\int_a^b f(-\tau)\,d\tau = \int_{-b}^{-a} f(\tau)\,d\tau,$$

so erhalten wir aus der letzten Gleichung

$$\int_0^\infty \Xi'(x, \xi, \tau+\sigma)\,d\tau = \int_{-\infty}^0 \Xi'(\xi, x, \tau-\sigma)\,d\tau. \tag{24}$$

Mit Worten heißt dies:

Die Verlustarbeit, welche durch die von der Zirkulation an der Stelle x herrührende Störungsgeschwindigkeit an der Stelle ξ verursacht wird, ist ebenso groß, wie die Verlustarbeit, welche an der Stelle x entstehen würde, wenn die von ξ ausgehenden Schraubenwirbel anstatt nach hinten nach

vorn in der Verlängerung der wirklichen Wirbel verlaufen würden.

Die gesamte durch die beiden Zirkulationselemente $\Gamma(x)\,dx$ und $\Gamma(\xi)\,d\xi$ herrührende Verlustarbeit ergibt sich nun nach Gleichung (19) und (24)

$$d^2 W = \varrho \frac{\Gamma(x)\cdot\Gamma(\xi)}{4\pi}\left[\int_0^\infty \Xi'(\xi, x, \tau-\sigma)\,d\tau + \int_{-\infty}^0 \Xi'(\xi, x, \tau-\sigma)\,d\tau\right] dx\cdot d\xi$$

$$= \varrho \frac{\Gamma(x)\cdot\Gamma(\xi)}{4\pi}\left[\int_{-\infty}^{+\infty} \Xi'(\xi, x, \tau-\sigma)\,d\tau\right] dx\cdot d\xi. \qquad (25)$$

Dies läßt sich folgendermaßen ausdrücken:

Die gesamte Verlustarbeit, welche durch die gegenseitige Einwirkung zweier Zirkulationselemente an den Stellen x und ξ dieser Elemente verursacht wird, ist ebenso groß, wie die Verlustarbeit, welche an der einen Stelle allein entstehen würde, wenn die von der anderen Stelle ausgehenden Schraubenwirbel sich nach beiden Seiten ins Unendliche erstrecken würden.

Wenn wir Gleichung (25) über jeden Flügel nach ξ integrieren und die Summe über die n Flügel bilden, so erhalten wir die Verlustarbeit dW, welche das Element $\Gamma(x)\,dx$ in Verbindung mit sämtlichen Schraubenflügeln verursacht. Bei der Integration und Summation der rechts stehenden Ausdrücke ist zu beachten, daß

$$\frac{1}{c(x)}\cdot\sum_1^n \int_0^R \frac{\Gamma(\xi)}{4\pi}\int_0^\infty \Xi'(\xi, x, \tau-\sigma)\,d\tau\,d\xi = w(x),$$

die zur Schraubenfläche senkrechte Störungsgeschwindigkeit im Punkte x und

$$\frac{1}{c(x)}\cdot\sum_1^n \int_0^R \frac{\Gamma(\xi)}{4\pi}\int_{-\infty}^{+\infty} \Xi'(\xi, x, \tau-\sigma)\,d\tau\,d\xi = w'(x)$$

die entsprechende Störungsgeschwindigkeit auf der durch x gehenden Schraubenlinie in großer Entfernung hinter der Schraube ist, wo die Schraubenwirbel nach beiden Seiten hin sich erstrecken. Wir erhalten demnach

$$dW = \varrho\cdot\Gamma(x)\cdot w'(x)\cdot c(x)\cdot dx. \qquad (26)$$

Der weitere Gang der Untersuchung ist derselbe wie beim ersten Beweis. Wir haben $w'(x)$ zu bestimmen aus der Bedingung:

$$\text{Verlustarbeit} = W = \varrho \sum_1^n \int_0^R \Gamma(x) \cdot w'(x) \cdot c(x) \cdot dx = \text{Minimum} \qquad (27)$$

$$\text{Schubarbeit} = S \cdot v = \varrho v \sum_1^n \int_0^R \Gamma(x) \cdot x \cdot \omega \cdot dx = \text{Konst.} \qquad (28)$$

was auf die Gleichung führt

$$\Gamma(x) \cdot w'(x) \cdot c(x) = C \cdot v \cdot \Gamma(x) \cdot x \cdot \omega, \qquad (29)$$

wobei C Lagrange'scher Faktor ist. Daraus ergibt sich

$$\frac{w'}{x\,\omega} = C \cdot \sin \varepsilon, \qquad (30)$$

wenn wir wieder $\frac{v}{c} = \sin \varepsilon$ einführen. Dieser Ausdruck stimmt mit der früher abgeleiteten Formel (3) überein und w' bedeutet auch hier wieder die Geschwindigkeit hinter der Schraubenfläche und ist identisch mit der durch Drehen einer starren Schraubenfläche erhaltenen Geschwindigkeit.

Wir haben durch die vorhergehende Untersuchung die Lösung des Minimalproblems der Schraubenpropeller in der Form gewonnen, daß wir angeben können, wie die Geschwindigkeitsverteilung in großer Entfernung hinter der Schraube sein muß. Für die Beurteilung der Vorgänge an den Schraubenflügeln interessiert vielfach die Störungsgeschwindigkeit, welche an der Stelle der Flügel selbst herrscht. Für Schrauben, deren Flügel symmetrisch zueinander liegen[1]), was ja fast stets zutrifft, läßt sich nun folgender einfacher Satz aufstellen:

Bei symmetrischen Schrauben sind die Störungsgeschwindigkeiten am Flügel halb so groß, wie an den entsprechenden Stellen der Schraubenfläche weit hinter dem Propeller.

Dieser Satz ist leicht einzusehen: Die von der Zirkulation um die einzelnen Flügel herrührenden Störungsgeschwindigkeiten heben sich am Ort eines Flügels infolge der symmetrischen Lage der in Frage kommenden Stellen zu den übrigen Flügeln gerade auf. Es ist also nur der Einfluß der von den Flügeln abgehenden Wirbelsysteme zu betrachten. Nun ist aber das Wirbelsystem von 0 bis $+\infty$, wie es an den Flügeln wirksam ist,

1) Dazu gehört, daß die Flügel mit Radien zusammenfallen, welche in ein und derselben Ebene senkrecht zur Achse (Schraubenkreisebene) liegen und gleichen Winkelabstand voneinander haben.

kongruent zu dem System von 0 bis $-\infty$, welches die gedachte Fortsetzung des ersteren bildet. Wegen der symmetrischen Lage eines jeden Punktes der Flügel zu den beiden Wirbelsystemen ist auch die von den beiden Systemen verursachte Störung dieselbe. Weit hinter der Schraube, wo ein System von $-\infty$ bis $+\infty$, also die Summe der beiden einander kongruenten Systeme wirksam ist, wird demnach auch die Störung doppelt so groß wie am Flügel selbst, wo nur das eine System in Wirksamkeit tritt. Formelmäßig können wir diesen Gedanken so ausdrücken: Ein Flügelelement an der Stelle x erzeugt an einer anderen Stelle ξ eines Flügels eine Störung, welche abgesehen von nebensächlichen Faktoren gegeben ist durch die Funktion

$$\int_0^\infty \Xi'(x, \xi, \tau + \sigma) d\tau,$$

wenn wieder σ den Winkel bezeichnet, den die durch x und ξ gehenden Radien bilden. Wegen der symmetrischen Anordnung ist ein dem Flügelelement x entsprechendes symmetrisch zu ξ gelegenes Element vorhanden, dessen Einfluß durch den Ausdruck

$$\int_0^\infty \Xi'(x, \xi, \tau - \sigma)\, d\tau$$

gegeben ist. Nehmen wir die von diesen beiden Elementen verursachten Störungen zusammen und berücksichtigen, daß wir $\tau + \sigma$ wegen des symmetrischen Baues von Ξ' mit $-\tau - \sigma$ und $\tau - \sigma$ mit $-\tau + \sigma$ vertauschen können. (Vergl. S. 209), so erhalten wir:

$$\begin{aligned} &\int_0^\infty \Xi'(x, \xi, \tau + \sigma)\, d\tau + \int_0^\infty \Xi'(x, \xi, \tau - \sigma)\, d\tau \\ &= \int_0^\infty \Xi'(x, \xi, -\tau - \sigma)\, d\tau + \int_0^\infty \Xi'(x, \xi, -\tau + \sigma)\, d\tau \\ &= \int_{-\infty}^0 \Xi'(x, \xi, \tau - \sigma)\, d\tau + \int_{-\infty}^0 \Xi'(x, \xi, \tau + \sigma)\, d\tau. \qquad (31) \end{aligned}$$

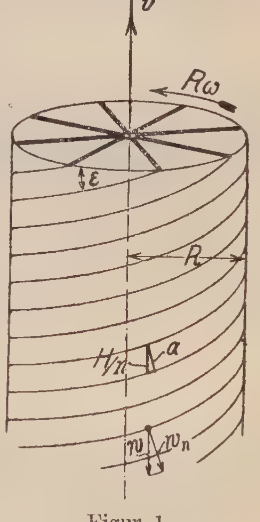
Hieraus ist die Gleichwertigkeit des Wirbelsystemes hinter dem Ausgangspunkt (von 0 bis $+\infty$) mit dem davor liegenden (von $-\infty$ bis 0) auch formelmäßig ersichtlich.

Zum Schluß möge noch etwas über die Analogie des hier behandelten Schraubenproblems mit dem in der Einleitung erwähnten Tragflügelproblem gesagt werden, das ja den Anstoß zu der ganzen Untersuchung gab: Dem Eindecker würde eine einflügelige Schraube

entsprechen. Der dort auftretenden Funktion $\frac{1}{x-\xi}$ entspricht hier die Funktion $\Psi(x, \xi, 0)$. Dem ungestaffelten Mehrdecker entspricht die Schraube mit symmetrischer Flügelanordnung; dem gestaffelten Mehrdecker die unsymmetrische Schraube. Der Vertikalabstand der einzelnen Decks spielt dort etwa dieselbe Rolle wie hier der Winkel σ.

Zusatz.

Die von Herrn Betz gefundene Vorschrift für die günstigste Kraftverteilung bei Luftschrauben erlaubt in dem Fall, daß es sich um eine vielflügelige Luftschraube handelt, auf ziemlich einfache Weise Näherungsformeln für die Verteilung der Zirkulation — und damit auch der Luftkräfte und der Flügelgestalt — aufzustellen. Entsprechend dem Gedankengang im Abschnitt 7 der II. Mitteilung meiner „Tragflügeltheorie"[1]) wird das System von Schraubenflächen, das durch die Luftschraube bei ihrer Bewegung in die Luft hineingeschnitten wird (Fig. 1), als starres Gebilde betrachtet, und es wird diesem starren Gebilde durch einen Stoß eine Geschwindigkeit w in der Achsenrichtung erteilt. Die bei diesem Stoß erzeugten Luftgeschwindigkeiten stimmen im Bereich der Theorie der I. Ordnung mit den Geschwindigkeiten überein, die die Schraube hinter sich in der Luft zurückläßt. In der äußeren Randzone treten beim Umströmen der Kanten der Schraubenflächen Radialgeschwindigkeiten auf, die eine besondere Betrachtung erfordern. Diese Radialgeschwindigkeiten klingen aber, wenn die Schraubenflächen hinreichend dicht stehen, nach dem Innern zu sehr rasch ab, so daß es hier ausreicht, die tangentiale Komponente w_t und die axiale Komponente w_a der Luftbewegung zu betrachten. Die Bewegung, die bei dem Stoß erzeugt wird, erfolgt im Innern, wie leicht zu sehen, normal zu den Schraubenflächen, mittelst derer der Stoß auf die



Figur 1.

---

1) Nachrichten der Ges. d. Wiss. 1919, S. 107.

Flüssigkeit ausgeübt wird. Ist v die Geschwindigkeit, mit der die Luftschraube gegen die Luft fortschreitet, ω deren Winkelgeschwindigkeit, dann ist die Steigung der Schraubenflächen $H = v\,T$, wo $T = 2\pi/\omega$ die Umlaufszeit ist, also

$$H = 2\pi v/\omega.$$

Der Steigungswinkel ergibt sich aus

$$\operatorname{tg}\varepsilon = H/2\pi r = v/r\omega,$$

wo r der Abstand des Elementes der Drehachse ist. Der Geschwindigkeit des Schraubenflächensystems nach dem Stoß w entspricht somit eine normal zu den Schraubenflächen gerichtete Flüssigkeitsbewegung vom Betrage $w_n = w\cos\varepsilon$; dies ergibt die Komponenten:

$$w_t = w\cos\varepsilon\sin\varepsilon = w\frac{v\,r\,\omega}{v^2+r^2 w^2} \qquad (1)$$

$$w_a = w\cos^2\varepsilon \qquad = w\frac{r^2\omega^2}{v^2+r^2\omega^2}. \qquad (2)$$

Die Zirkulation für einen Schraubenflügel im Radius r ergibt sich bei n Schraubenflügeln, die gleichmäßig auf den Umfang verteilt sein mögen, zu

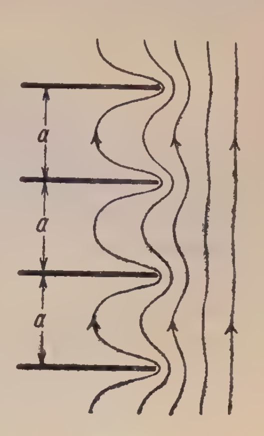
$$\Gamma = \frac{2\pi r w_t}{n} = \frac{2\pi w}{n}\cdot\frac{v\,r^2\omega}{v^2+r^2\omega^2} \quad \left(= \frac{h\cdot w_a}{n}\right). \qquad (3)$$

Anmerkung: Die Ergebnisse dieser Betrachtung, die ebenso wie die Betz'sche Theorie nur für den Grenzfall sehr schwach belasteter Schrauben streng gültig sind, lassen sich auf stark belasteten Schrauben erweitern, wenn man, wie Herr Betz in einer anderen, noch zu veröffentlichenden Untersuchung gezeigt hat, in den obigen Formeln für v und $r\omega$ die genauen Werte der Strömungsgeschwindigkeiten relativ zu den Flügelprofilen $v + w_a/2$ und $r\omega - w_t/2$ setzt, wobei unter v die Zuflußgeschwindigkeit im Unendlichen vor der Schraube und unter w_a und w_t die zusätzlichen Geschwindigkeitskomponenten weit hinter der Schraube verstanden werden. Eine brauchbare Näherung für mäßig belastete Schrauben erhält man übrigens schon dadurch, daß in den obigen Formeln v durch $v + w/2$ ersetzt wird, wobei die einfache Vorstellung von in die Luft geschnittenen Schraubenflächen, die sonst hinfällig wird, beibehalten werden kann. Von dieser letzteren Näherung ist in Formel (8) Gebrauch gemacht.

Für die Randzone gilt die folgende Sonderbetrachtung. Es sei a der senkrechte Abstand von zwei benachbarten Randkurven der Schraubenflächen, also gemäß Figur 1

$$a = \frac{H}{n} \cdot \cos \varepsilon = \frac{2\pi v r}{n\sqrt{v^2 + r^2\omega^2}}. \tag{4}$$

An die Stelle der Umströmung der Ränder der Schraubenflächen mag nun die Umströmung der Ränder eines äquidistanten Ebenensystems vom Abstand a, gemäß Fig. 2 gesetzt werden, um den Abfall der Zirkulation an den Flügelspitzen zu studieren. Nach den Darlegungen von Nr. 6 der II. Mitteilung ist die Zirkulation gleich dem Potentialsprung der in der Flüssigkeit zurückgelassenen Strömung zu setzen. Es ist also das Strömungspotential für die in Fig. 2 angedeutete Strömung zu berechnen. Die nach Fig. 2 zerschnittene Ebene (z-Ebene) läßt sich konform auf das Äußere des Einheitskreises (t-Ebene) abbilden durch



Figur 2.

$$z = \frac{a}{\pi} \ln \frac{1}{2}\left(t + \frac{1}{t}\right). \tag{5}$$

Die Strömung von Fig. 2 verwandelt sich dabei in zirkulatorisches Umströmen des Einheitskreises, es ist also $\Phi + i\Psi = F = ic\ln t$; hiermit ergibt sich nach kurzer Rechnung:

$$z = \frac{a}{\pi} \ln \cos \frac{F}{c}. \tag{6}$$

Für reelle Werte von F, also für die Stromlinien $\Psi = 0$, wird

$$x = \frac{a}{\pi} \ln \cos \frac{\Phi}{c} \quad \text{oder} \quad \Phi = \pm c \arccos e^{\pi x/a}$$

für $x \gtreqless 0$. Der Potentialsprung wird somit, wenn, mit $R =$ Außenradius, $R - r$ für $-x$ gesetzt wird

$$[\Phi] = 2c \arccos e^{-\pi(R-r)/a}. \tag{7}$$

Der Abfall der Zirkulation an den Flügelspitzen der Schraube wird nun proportional dem Abfall des Potentialsprunges in Gleichung (7) angenommen. Bei den praktisch üblichen Abmessungen liegen die Funktionswerte von Γ nach Gleichung (3) für die Gegend der Flügelspitzen regelmäßig schon in der Nähe des Asymptotenwertes, sind also dort fast konstant. Es wird deshalb kein großer Fehler gegenüber der wahren Lösung der Minimumsaufgabe entstehen, wenn man an ihrer Stelle denjenigen Verlauf von Γ setzt, der sich ergibt, wenn man die Werte von Γ aus Gleichung (3)

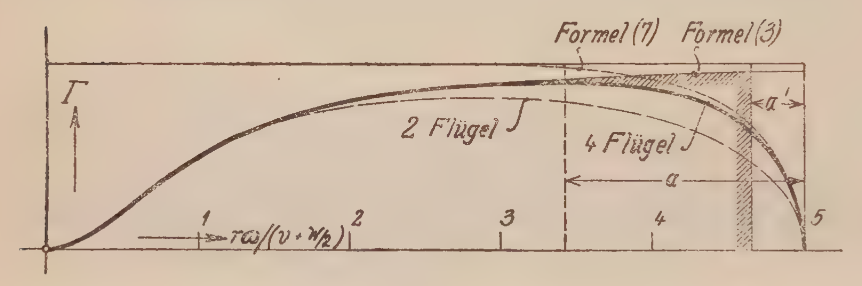
mit dem Ausdruck der Gleichung (7) multipliziert, wobei durch die Wahl $c = 1/\pi$ der Asymptotenwert dieses Ausdruckes auf 1 normiert wird. Hiermit ergibt sich also als Endformel:

$$\Gamma = \frac{4w}{n} \cdot \frac{(v+w/2)\,r^2\omega}{(v+w/2)^2 + r^2\omega^2} \cdot \arccos e^{-\pi(R-r)/a} \tag{8}$$

mit a nach Gleichung (4), in der ebenfalls gemäß der Anmerkung S. 214 v durch $v+w/2$ ersetzt werden kann. Die Gleichung gilt umso genauer, je mehr Schraubenflügel vorhanden sind, es ist aber zu erwarten, daß sie auch bei wenigen, ja selbst bei zwei Flügeln, immer noch brauchbare Resultate ergibt. Die Fig. 3 zeigt in der stark ausgezogenen Linie den Verlauf von Γ abhängig von r für vier Flügel und für einen Wert

$$H/2\pi R = (v+w/2)/R\omega = 1:5,$$

was etwa dem bei Flugzeugschrauben Üblichen entspricht. Der Verlauf von Γ bei zwei Flügeln ist zum Vergleich in Fig. 3 gestrichelt beigefügt.



Figur 3.

Bei Vergrößerung der Flügelzahl würde die Größe a verkleinert, der Abfall an der Flügelspitze also steiler werden; im Grenzfall von unendlich vielen Flügeln wäre dieser Abfall also senkrecht anzusetzen. Der Verlust an Völligkeit in der Verteilung Γ und daher auch in der Schubverteilung, der bei endlicher Flügelzahl gegenüber unendlich vielen Flügeln eintritt, läßt sich dadurch zum Ausdruck bringen, daß man die Verteilung nach Gleichung (7) flächengleich durch ein Rechteck ausgleicht. Dadurch erweist sich irgend eine Schraube als dynamisch gleichwertig mit einer Schraube mit unendlich vielen Flügeln, bei der die Flügel um einen Betrag a' (Figur 3) kürzer sind, als die Flügel der gegebenen Schraube. Der Betrag a' ergibt sich aus der Forderung, daß

$$\int_0^{c\pi/2} x\, d\Phi = \frac{c\pi}{2} \cdot a'$$

sein muß, zu

$$a' = \frac{a}{\pi} \ln 2 = 0{,}2207\, a. \tag{9}$$

In dem obigen Beispiel ergibt sich für unendlich viele Flügel (Verteilung in Fig. 3 durch Schraffur angedeutet) eine äquivalente Schraubenkreisfläche, die um 13 % kleiner ist, als die gegebene Schraubenkreisfläche der vierflügeligen Schraube.

Die Flügelkräfte und die erforderliche Formgebung der Flügel können, sobald Γ bekannt ist, nach den in der „Tragflügeltheorie" I 13 angegebenen Grundsätzen, die sich ohne weiteres auf den Fall der Schraube übertragen lassen, ermittelt werden. Ein näheres Eingehen hierauf kann deshalb an dieser Stelle unterbleiben.

L. Prandtl.

Die Entstehung des muscheligen Bruches.

Von

G. Tammann.

Vorgelegt in der Sitzung vom 16. Mai 1919.

Einen Riß in einer Glasplatte kann man mittelst eines heißen Körpers beliebig führen, indem man eine Stelle der Platte durch Berührung mit einem heißen Körper erwärmt. Der Riß strebt immer der erhitzten Stelle zu, in der der Druck erhöht ist. Wenn durch einen Schlag in einem Glasstück longitudinale Wellen erzeugt werden, und sich hierbei eine Rißfläche bildet, so wird diese von den Stellen erhöhten Druckes der Wellenfläche abgelenkt werden. Dadurch entstehen auf der Rißfläche konzentrisch um die Schlagstelle verlaufende Wellentäler und Wellenberge, deren Tiefe und Höhe von der Schlagstelle aus abnehmen können, wenn die Dämpfung der longitudinalen Welle eine merkliche ist. Wäre die Bruchfläche an ihrem Ursprung eine Ebene, so könnte man durch Zählung der Wellentäler auf der einen Bruchfläche bestimmen, ob ihnen Verdichtungen oder Verdünnungen der Longitudinalwelle entsprechen, da die erste halbe Welle einer Verdichtung entsprechen sollte. Aber gerade die Ursprungstelle der Bruchfläche ist in der Regel so wenig regelmäßig, daß eine solche Bestimmung nicht ausführbar ist. Daß beim Zerschlagen eines Stückes nicht nur longitudinale Schwingungen einer Wellenlänge auftreten, ersieht man auf der Bruchfläche, auf der man häufig mehrere Wellensysteme von einer Wellenlänge findet, die von 100 bis 1 mm schwanken kann.

Gleichgültig ob der Körper ein homogenes Glas, ein einzelner Kristall oder ein Kristalliten-Konglomerat ist, seine Bruchfläche wird muschelig sein, wenn nur den Kristalliten die Spaltbarkeit

und die Fähigkeit Gleitebenen zu bilden fehlt und die Kristallite gut miteinander verklebt sind. In der Tat findet sich unter diesen Bedingungen auch bei Kristalliten-Konglomeraten ein ausgeprägter muscheliger Bruch, wie beim aluminothermisch hergestellten Mangan und vielen Metallverbindungen, denen sowohl Spaltbarkeit als auch Gleitebenen-Bildung abgeht. Da die Spaltbarkeit der Kristalle eine sehr verschiedene ist, so wird man in Kristalliten-Konglomeraten alle Übergänge vom muscheligen bis zum körnigen Bruch beobachten können.

Nicht selten treten auf den muscheligen Bruchflächen noch Strahlen auf, die von der Schlagstelle aus radial, also senkrecht zu Wellenbergen und Wellentälern auf den Bruchflächen verlaufen. Auf der einen Bruchfläche sind diese Strahlen Vertiefungen, auf der anderen Erhöhungen. Die Ursache ihrer Bildung ist wohl in einer Störung der Longitudinalwelle zu suchen, die durch Inhomogenitäten oder andere Ursachen hervorgerufen werden kann.

Ueber den Einfluss des Druckes auf die Explosionstemperatur.

Von

G. Tammann.

Vorgelegt in der Sitzung vom 20. Juni 1919.

Der Explosivstoff wurde in einem kleinen, an einem Ende offenen Glasröhrchen in einen mit Quecksilber gefüllten Stahlzylinder gebracht. Dieser war mit einem Federmanometer und einer Hochdruckpumpe verbunden. Damit das Öl aus der Hochdruckpumpe nicht in den Stahlzylinder mit dem Explosivstoff gelangen konnte, war zwischen beide ein mit Quecksilber gefülltes U-Rohr geschaltet. Der Stahlzylinder mit dem Explosivstoff tauchte in ein Ölbad, welches durch einen rotierenden Rührer kräftig umgerührt wurde. Die Temperatur des Bades und der Druck im Stahlzylinder wurden von Minute zu Minute beobachtet, während das Bad erwärmt wurde. Bei Eintritt der Explosion sprang der Zeiger des Manometers plötzlich zu einem um 100 bis 300 kg/qcm höheren Druck, wobei der ganze Apparat erzitterte. Darauf fiel der Manometerzeiger schnell, um schließlich nach 1 bis 2 Minuten wieder langsam, wie vor der Explosion infolge der Erwärmung des Zylinders zu steigen.

Die Mengen des Explosivstoffs mußten natürlich so gewählt werden, daß die plötzlichen Drucksteigerungen nicht zu groß waren, sich ungefähr in jenen Grenzen bewegten.

In den folgenden Tabellen sind angegeben:

1) die Erhitzungsgeschwindigkeit in Graden pro 10 Minuten, vor Erreichung der Explosionstemperatur;
2) die Mengen des Explosivstoffs in gr oder ccm;

3) und 4) der Druck und die Temperatur, bei denen die Explosion eintrat;

5) die Differenz Δp der Drucke vor und nach der Explosion; diese Werte schwanken stark in Folge verschiedenen Funktionierens der Übertragung von der Manometerfeder zum Zeiger, so daß sogar negative Δp-Werte zur Beobachtung kamen;

6) der plötzliche Druckanstieg während der Explosion. Auch diese Werte ändern sich aus demselben Grunde nicht regelmäßig mit der Menge des Explosivstoffes und dem Druck.

Versuchs-Nr.	Erhitzungsgeschwindigkeit	Menge des Nitroglycerins in gr	Explosions-		Mittel	Δp	Plötzliche Drucksteigerung
			Druck p kg/qcm	Temperatur			
			Nitroglycerin feucht.				
1	16,9°	0,079	188	155,5°	150,6	+122	+452
2	19,0	0,10	246	143,9	± 4,6	+ 69	+394
3	14,7	0,060	260	153,0		+ 60	+330
4	9,4	0,211	935	121,5	134,5	+153	+455
5	19,0	0,075	1051	136,0	± 8,6	+ 21	+ 69
6	9,5	0,070	1102	146,0		+ 36	+123
7	9,6	0,101	1622	144,2		— 3	+ 29
8	9,7	0,101	2011	157,0	152,0	leck	+ 30
			leck		± 5,2		
9	8,6	0,067	2130	154,8		+ 14	+ 62
10	14,3	0,072	2667	141,0	141,9	+ 25	+118
11	14,0	0,075	2968	142,9	± 1	— 8	+ 82
			Nitroglycerin trocken.				
1	9,8	0,086	165	155,0		+ 95	+460
2	9,3	0,090	1112	158,0		+ 35	+185
3	8,8	0,086	1160	152,0		+ 30	+250
4	9,7	0,082	2421	151,0		— 39	+ 49
5	7,6	0,086	2493	153,3		leck	+ 57

Beim Öffnen des Stahlzylinders entwichen die bei der Explosion gebildeten Gase; außer den gasförmigen war die Bildung fester oder flüssiger Zersetzungsprodukte nicht festzustellen.

Die Explosionstemperaturen des feuchten Nitroglycerins schwanken ziemlich stark bei demselben Druck, nämlich um 5°, 9° und 5°. Eine Drucksteigerung um 2500 kg/cm² scheint die Explosionstemperatur um 8,7° zu erniedrigen, doch liegt dieser Einfluß noch an den Grenzen der Schwankungen. Bei trockenem Nitroglycerin sind diese Schwankungen geringer, aber auch hier

ist ein deutlicher Einfluß des Druckes auf die Explosionstemperatur nicht ersichtlich. Die Differenzen der Drucke vor und nach der Explosion, Δp, nehmen mit wachsendem Druck ab und sind schon bei 1500 kg/qcm innerhalb der Fehlergrenzen der Druckmessung nicht mehr merklich, d. h. die Volumina des Explosivstoffs und der gasförmigen Explosionsprodukte sind nicht mehr sehr verschieden voneinander. Auch die plötzliche Drucksteigerung während der Explosion nimmt mit wachsendem Drucke sehr schnell ab. Diese Drucksteigerung hängt aber in hohem Maße von der Einrichtung des Apparates ab. Bei diesen Versuchen war das Manometer von dem Explosionsgefäß durch 2 m lange Stahlkapillaren mit einer lichten Weite von ·2 mm getrennt.

Tetranitromethan und Benzol.
(2 Mol. + 1 Mol.)

Versuchs-Nr.	Erhitzungsgeschwindigkeit	Menge der Mischung in ccm	Explosions-		Δp	Plötzliche Drucksteigerung
			Druck kg/qcm	Temperatur		
1	15,3°	0,105	380	163,3°	+ 50	+170
2	19,9	0,112	1100	171,0	0	+150
3	14,0	0,094	2160	162,0	leck	?

Hier ist noch weniger wie beim Nitroglycerin ein Druckeinfluß auf die Explosionstemperatur mit Sicherheit festzustellen. Beim Öffnen des Stahlzylinders nach der Explosion entwich kein Gas. Die Röhrchen enthielten ein wenig einer öligen Flüssigkeit. Es scheint sich bei der Explosion ein Kondensationsprodukt aus Tetranitromethan und Benzol gebildet zu haben; daher sind auch die Δp-Werte hier nicht deutlich positiv. Die untersuchte Mischung von Tetranitromethan und Benzol verliert mit der Zeit, im Laufe von drei Tagen, ihre Explosionsfähigkeit. Bei zwei Versuchen, die drei Tage nach Herstellung der untersuchten Mischung vorgenommen wurden, erfolgte beim Erhitzen auf 210° bei den Drucken von 130 bis 350 kg/qcm und beim Erhitzen bis 205° unter Drucken von 2700 bis 2945 kg/qcm keine Explosion. Es scheint sich also in dem explosiven Gemisch bei Zimmertemperatur eine langsame Reaktion zu vollziehen, durch die jene Mischung die Explosionsfähigkeit verliert.

Tetranitromethan und Naphtalin.
(2 Mol. + 1 Mol.)

Versuchs-Nr.	Erhitzungsgeschwindigkeit	Menge der Mischung in ccm	Explosions- Druck kg/qcm	Explosions- Temperatur	Δp	Plötzliche Drucksteigerung
1	19,4°	0,082	258	160,4	+137	+282
2	13,3	0,020	952	164,2	+143	+188
3	11,5	0,022	2290	166,1	+ 13	+108

Hier steigt bei einer Drucksteigerung von ca. 2000 kg/qcm die Explosionstemperatur um 5,7° an. Da wegen der geringen Anzahl der Versuche die Schwankungen der Explosionstemperatur bei demselben Drucke nicht bekannt sind, so ist es sehr wohl möglich, daß auch hier dieser Druckeinfluß noch innerhalb der Schwankungen liegt, denen die Temperatur der Explosion bei verschiedenen, sonst aber unter gleichen Umständen ausgeführten Versuchen, unterliegt. Bei den Versuchen mit diesen Mischungen entwichen beim Öffnen des Stahlzylinders die Explosionsgase.

Die erheblichen Schwankungen der Explosionstemperatur haben die Feststellung des Druckeinflusses auf dieselbe sehr ungenau gemacht. Man kann sagen, daß der Einfluß einer Drucksteigerung um 2500 kg/cm² auf die Explosionstemperatur nicht größer ist als ihre Schwankungen, die im Mittel etwa 9° betragen.

Bei der Explosionstemperatur wird die Reaktionsgeschwindigkeit, die zur Explosion führt, stürmisch, weil bei dieser Temperatur die Reaktionswärme hinreicht, um unter den gegebenen Verhältnissen der Wärmeableitung die Reaktionsgeschwindigkeit beschleunigt zu steigern. Die Schwankungen der Explosionstemperatur können verursacht werden durch atomistische Verhältnisse, indem in gewissen Bezirken unwahrscheinlich hohe lebendige Kräfte mehrerer zusammenstoßender Moleküle oder andere unwahrscheinliche atomistische Ereignisse auftreten, die zu einer abnorm hohen Reaktionsgeschwindigkeit führen, die ihrerseits wieder abnorme Wärmequantitäten entwickelt. Es können aber auch andere Ursachen die Schwankungen bedingen, wie nicht gleichmäßig verteilte geringe Beimengungen, die schon vor dem Beginn der Reaktion vorhanden sind oder sich bei ihr bilden, oder die an den Wänden der Glasröhrchen haften.

In der Tat waren beim feuchten Nitroglycerin, das durch Wassertröpfchen getrübt war, die Schwankungen sehr viel größer als beim trockenen Nitroglycerin. Um den Druckeinfluß auf die Explosionstemperatur genauer zu bestimmen, müßten vor allem auf Grund eines genaueren Studiums der Schwankungen diese zu verkleinern oder zu beseitigen sein.

Der Druckeinfluß auf die Explosionstemperatur flüssiger Explosivstoffe wird bestimmt durch den Druckeinfluß auf das Wärmeleitvermögen des Explosionsstoffes und seiner Umgebung, und dann durch Änderung der chemisch-kinetischen Bedingungen, wie die Vergrößerung der Zahl molekularer Zusammenstöße bei isothermer Drucksteigerung infolge der Volumenverkleinerung des Sprengstoffes. Wenn diese beiden Faktoren allein den Druckeinfluß bestimmen, so ist ein geringer Druckeinfluß auf die Explosionstemperatur zu erwarten.

Die Schlagempfindlichkeit flüssiger Explosivstoffe kann also nicht als eine Folge der den Schlag begleitenden Druckerhöhung aufgefaßt werden, auch ist die Überschreitung der Explosionstemperatur infolge adiabatischer Temperatursteigerung bei flüssigen und festen Stoffen unwahrscheinlich. Daher ist es wahrscheinlich, daß neben der Größe des bei dem Schlage erzeugten Drucks die Geschwindigkeit seines Anwachsens von ausschlaggebender Bedeutung ist. Hiermit scheinen auch die anderen Tatsachen, vor allem die Zündung durch Influenz, in Einklang zu stehen.

Ueber Anlauffarben von Metallen.

Von

G. Tammann.

Vorgelegt in der Sitzung vom 16. Mai 1919.

Bekanntlich treten beim Erhitzen von Metallen an der Luft Anlauffarben auf. Die Reihenfolge dieser Farben ist im allgemeinen die, welche Luftschichten wachsender Dicke zeigen. Die Metalle überziehen sich also, so weit sie Anlauffarben zeigen, mit dünnen Oxydhäuten, deren Verdickungsgeschwindigkeit aus der Änderung der Anlauffarben mit der Zeit bestimmt werden kann. Einige Metalle zeigen auch im Joddampf Anlauffarben und zwar schon bei gewöhnlicher Temperatur. Zur Auffindung des Gesetzes der Verdickungsgeschwindigkeit empfiehlt es sich aus experimentellen Gründen gerade diesen Vorgang zu untersuchen.

Es ergab sich, daß die isotherme Verdickungsgeschwindigkeit nur von dem Diffusionskoeffizienten des Jods in der das Ag bedeckenden AgJ-Schicht abhängt. Für den isothermen Vorgang folgt hieraus eine parabolische Abhängigkeit der Dicke von der Einwirkungszeit. Die im folgenden entwickelte und an der Erfahrung bestätigte Theorie der Anlauffarben hat mit der von Nernst entwickelten Kinetik des Auflösungsvorganges den Hauptgedanken gemeinsam. Der eigentliche chemische Vorgang verläuft sehr viel schneller als der Diffusionsvorgang, der in dem einen Fall das Agens an das Metall durch eine sich stetig verdickende Schicht der Reaktionsprodukte führt, und der im anderen Fall die Reaktionsprodukte von der Oberfläche des sich auflösenden Stoffes durch eine Schicht unveränderlicher Dicke fortführt. Die quantitativen Gesetze, welche die Anlauf- und die Auflösungsgeschwindigkeit beherrschen, sind entsprechend dem verschiedenen Verhalten der Schicht der Reaktionsprodukte verschieden.

Ueber die Einwirkung von Joddampf auf Metalle.

Bringt man in einen mit H_2SO_4 beschickten Exsikkator einige Jodkristalle und folgende Metalle, so beobachtet man eine Einwirkung des Joddampfes, die zur Entstehung von Oberflächenfarben führt, nur beim Ag, Cu, Pb und Tl. Auf diesen Metallen bilden sich mehr oder weniger gleichmäßige Schichten ihrer Jodverbindungen, die sich mit der Zeit verdicken und hierbei die bekannte Farbenskala dünner Schichten durchlaufen.

Bi und Sb werden bald dunkel gefärbt und auf ihrer Oberfläche bildet sich ein feiner dunkler Staub. Ähnlich verhält sich das Sn, auf dem sich eine gelbe, wenig zusammenhängende Schicht bildet. Es scheinen hier die Reaktionsprodukte sich zu submikroskopischen Kriställchen zu sammeln, wodurch die Entstehung einer auf der metallischen Fläche haftenden Schicht und damit die Entstehung von Oberflächenfarben verhindert wird.

Auf Fe, Co, Ni und Mn entstehen nach wenigen Minuten Pünktchen, die außerhalb des Exsikkators schnell zu Tröpfchen anwachsen, weil die entstehenden Jodide den Wasserdampf anziehen.

Gleichmäßiger, aber sehr viel langsamer, ist die Einwirkung auf Zn und Cd; auf Cr ist sie zweifelhaft. Nach mehreren Tagen ist das Zn wenig verändert, es ist ein wenig bläulicher geworden und durch schwache Ätzung sind die Umgrenzungen seiner Kristallite deutlich hervorgetreten; an Wasser gibt es merkliche Mengen von Jodzink ab. Beim Cd ist eine Veränderung nicht sichtbar, doch gibt es Spuren seines Jodides an Wasser ab. Poliertes Cr verliert seinen Glanz auch im Laufe einer Woche nicht.

Bei den Metallen, auf denen sich Oberflächenfarben bilden, kann die Geschwindigkeit der Verdickung der die Farben verursachenden Schicht bestimmt werden, wenn man für eine unveränderliche Konzentration des Joddampfes Sorge trägt, indem man die Änderung der Oberflächenfarbe in Abhängigkeit von der Zeit bestimmt.

Schon Newton hat die Dicke und Farbe von Luftschichten zwischen Glasplatten angegeben. Eine Darstellung dieser Theorie findet sich im Lehrbuch der Physik von Müller-Pouillet Bd. 2, S. 904, 1897. Von Rollet[1]) ist dieser Zusammenhang genauer fest-

---

1) Wiener Sitzungsberichte 77, III, S. 229, 1878 und Landolts Tabellen 1905, Tab. 186, S. 610.

gestellt worden, und da die Brechungskoeffizienten des Ag-J in Abhängigkeit von der Wellenlänge bekannt sind, so können auch für AgJ-Schichten verschiedener Dicke ihre Farben angegeben werden.

Zur Herstellung einer bestimmten unveränderlichen Jodkonzentration über einem Ag-Plättchen könnte man Luft über Jodkristalle leiten, um sie mit Joddampf zu sättigen. Die bei Zimmertemperatur noch nicht gesättigte Luft wirkt aber so schnell auf ein Ag-Plättchen ein, daß die Beobachtung der Änderung der Oberflächenfarben Schwierigkeiten macht. Leitet man Luft durch eine Lösung, die im Liter 0,002 g-Mol. J_2 und 0,014 g-Mol. KJ enthält, so ist die Jodkonzentration in der Luft erheblich kleiner und die Änderung der Anlauffarben kann bequem verfolgt werden. Da es bei den folgenden Versuchen nur darauf ankam, bei derselben J-Konzentration, deren Betrag nicht näher bekannt war, Vergleiche anzustellen, so brauchte für vollständige Sättigung der Luft nicht Sorge getragen zu werden. Ein Luftstrom, der von einer Wasserstrahlpumpe angesogen wurde, strich durch dieselbe Menge jener, bei jedem Versuch erneuerten Lösung und wurde über ein Ag-Plättchen von 1 cm^2 oberer Fläche geleitet, das sich in einem Chlorkalciumröhrchen befand. Die Farben traten auf dem Plättchen nicht gleichmäßig hervor. Der dem Luftstrom zugekehrte Teil des Plättchens zeigte beständig Farben größerer Dicke, als der dem Luftstrom abgewandte Teil. Beobachtet wurde der mittlere Teil des Plättchens, indem bei der Farbenbestimmung senkrecht auf das Plättchen gesehen wurde. Dabei hielt der Beobachter die Tabelle von Rollet in der Hand, um nach Ablauf je einer Minute die richtige Farbe oder die ihr entsprechende Dicke einer Luftschicht zu notieren. Die Farbenbezeichnungen der Tabelle erleichtern sehr die Wahl des Farbennamens, dasselbe gilt für die Verschiedenheit der Farben am vorderen und hinteren Ende des Plättchens.

Im Folgenden sind die Resultate von drei Versuchen bei 17° angegeben. Beim ersten Versuch sind angegeben die Farben der Mitte eines Ag-Plättchens in hartem Zustande, das durch Reiben mit einem Tuch kurz vor dem Versuch nochmals gereinigt wurde, für bestimmte Zeiten nebst den jenen Farben entsprechenden Dicken einer Luft- und einer AgJ-Schicht. Die Dicke der AgJ-Schicht ergibt sich aus der Luftschicht durch Multiplikation mit 0,445, weil nach Wernicke der Brechungsexponent des AgJ gleich 2,27 ist.

Die beiden folgenden Versuche wurden mit Ag-Plättchen von

anderer Oberflächenbeschaffenheit ausgeführt. Erhitzt man ein Silberband von etwa 0,5 mm Dicke im Bunsenbrenner bis zum Beginn seines Schmelzens, so wird, wie das Schmelzen beginnt, der am stärksten erhitzte Teil plötzlich viel blanker als der nicht oder wenig erhitzte Teil. Zwischen beiden hat sich ein breiter Streifen mit auffallend blinder, durch die Rekristallisation veränderter Oberfläche gebildet. Man kann also einem Ag-Streifen leicht drei verschiedene Oberflächen erteilen, 1) die ursprüngliche nach dem Walzen gereinigte, 2) die ganz blanke, geschmolzen gewesene mit großen Kristalliten und 3) die ganz rauhe mit vielen kleinen dendritischen Kriställchen.

1) Gereinigte Walzoberfläche.

Zeit in Minuten	Farbe	Dicke der Luftschicht in $\mu\mu$	Dicke der AgJ-Schicht in $\mu\mu$ gefunden	berechnet
1	indigo	282	124	125
2	gelbgrün	409	178	176
3	rot	490	216	216
4	blau	570	247	249
5	grün	650	286	279
6	gelbgrün	680	299	305
7	falbes gelb	720	319	330
8	purpur matt	780	335	353
9	meergrün	870	383	374
12	graurot	996	440	432
14	blaugrün	1168	515	489
17	fleischrot	1264	554	514
21	blaugrün	1450	638	571

2) Ganz blanke Schmelzoberfläche.

Zeit in Minuten	Dicke der AgJ-Schicht in $\mu\mu$ gefunden	berechnet
0,25	59	73
0,50	101	104
1	154	147
2	205	207
3	229	254
4	286	293
5	330	328
6	402	359
7	440	389

3) Ganz rauhe Oberfläche.

Zeit in Minuten	Dicke der AgJ-Schicht in $\mu\mu$ gefunden	berechnet
0,5	62	17
1	72	25
2	101	111
3	124	136
4	154	157
5	191	176
6	205	192
7	215	208
8	229	222
9	242	236
10	264	249
11	286	261
12	299	272

Die Verdickungsgeschwindigkeit der AgJ-Schicht auf der reinsten und glattesten, nämlich der Schmelzoberfläche, ist, wie zu erwarten war, am größten, ihr folgt die der AgJ-Schicht auf der gereinigten Walzoberfläche, welche nicht viel unebener als die Schmelzoberfläche ist, aber sicher mehr Fremdstoffe als diese enthält, und am kleinsten ist die Verdickungsgeschwindigkeit auf der reinen aber rauhen, rekristallisierten Oberfläche. Durch Aufrauhen der Oberfläche wird naturgemäß die Zeit, die zur Herstellung einer AgJ-Schicht von bestimmter Dicke bei gleicher Konzentration des Joddampfes in der Luft und bei gleicher Strömungsgeschwindigkeit der Luft nötig ist, vergrößert.

Auf der Schmelzoberfläche ist besonders bei größeren Dicken der AgJ-Schicht die verschiedene Färbung einzelner Gruppen von Kristalliten zu erkennen. Ob diese Verschiedenheit der Färbung auf eine verschiedene Angriffsgeschwindigkeit verschieden orientierter Kristallflächen oder auf eine verschiedene Neigung derselben zum Luftstrom, also auf eine Verschiedenheit der Strömungswirbel, zurückzuführen ist, mag dahin gestellt sein. Die Feststellung der allgemeinen Farbe eines größeren Bezirkes der Oberfläche, wird durch ganz andere Färbung einzelner Kristallite nicht erheblich gestört. Am leuchtendsten treten die Farben auf der Schmelzoberfläche hervor und am mattesten auf dem rauhen Plättchen.

Die Anlauferscheinungen beim Cu (hart mit gereinigter Oberfläche) sind denen beim Ag mit derselben Oberflächenbeschaffenheit ganz ähnlich. Da der Brechungsexponent des CuJ nicht bekannt ist, so können die Dicken der den beobachteten Farben ent-

sprechenden Luftschichten nicht auf die Dicken der CuJ-Schichten reduziert werden. Nach gleichen Zeiten sind jene Dicken der Luftschichten um 10 Proz. größer als die bei Ag. Für den Verlauf der Verdickung gilt das beim Ag noch anzugebende Gesetz. Sehr viel geringer als beim Ag und Cu ist die Verdickungsgeschwindigkeit der Oberflächenschicht bei der Einwirkung von Joddampf auf Blei. Die Strömungsgeschwindigkeit der Luft, ihr Jod-Gehalt und ihre Temperatur waren dieselben wie bei den Versuchen mit Ag und Cu. Da der Brechungsexponent des PbJ nicht bekannt ist, so sind die den Anlauffarben entsprechenden Dicken einer Luftschicht angegeben. Die halben Werte dieser Dicken werden angenähert den Dicken des PbJ_2 entsprechen. Die eine Seite des Bleiplättchens war unmittelbar vor dem Versuch frisch geschabt und betrug 1 cm^2. 2p beträgt hier nur $6{,}4 . 10^{-10}$ mm und würde bei der Umrechnung von Luft auf PbJ_2-Dicken auf etwa $^1/_4$ dieses Wertes abnehmen.

Zeit in Minuten	Dicke der Luftschicht, die der Anlauffarbe entspricht in $\mu\mu$	
	gefunden	berechnet
28	140	134
41	164	162
85	235	233
100	245	253
117	257	274
130	282	288
170	352	330

Die Dicke y der Anlaufschicht in Abhängigkeit von der Einwirkungszeit t wird durch ein einfaches Gesetz bestimmt. Die Verdickungsgeschwindigkeit $\frac{dy}{dt}$ wird durch die Menge Jod, die zur Oberfläche des Ag gelangt, bestimmt sein, und da die Jodkonzentration an der Oberfläche der AgJ-Schicht durch die Strömungswirbel unveränderlich erhalten wird, wird jene Jodmenge umgekehrt proportional der Schichtdicke y des AgJ sein. Also

$$\frac{dy}{dt} = \frac{p}{y}.$$

Dementsprechend sollte für y und t die Beziehung

$$y^2 = 2pt$$

bestehen. In der Tat stellt diese Gleichung die Beziehung zwischen y und t innerhalb der Fehlergrenzen der Beobachtung dar. Die

Dicken der Anlaufschicht liegen also auf dem positiven Ast einer Parabel, deren Scheitel mit dem Nullpunkt der $y-t$-Koordinaten zusammenfällt, und deren Symmetrieachse die t-Achse ist.

Berechnet man für die drei Versuchsreihen am Ag aus jeder Beobachtung den Wert $2p$, nimmt das Mittel aus diesen Werten und berechnet nun für vorgelegte Werte von t die von y, so erhält man die in den Tabellen 1) 2) und 3) als berechnet bezeichneten Dicken. Die Mittelwerte von $2p$ sind für den Versuch 1) $155{,}5 . 10^{-10}$ mm, für den Versuch 2) $215 . 10^{-10}$ und für den Versuch 3) $62 . 10^{-10}$.

Die Uebereinstimmung zwischen den gefundenen und berechneten Dicken ist bei den Versuchsreihen 1) und 2) bis auf die größten Schichtdicken ganz befriedigend. Bei diesen ändert sich die Dicke mit der Farbe viel mehr, als bei den dünneren Schichten (siehe Rollets Tabelle), wodurch die Dickenbestimmung unsicher wird. Bei Versuch 3) ist die Übereinstimmung am schlechtesten; wegen der Rauheit der Oberfläche sind hier die Verhältnisse am unübersichtlichsten.

Die Konstante $2p$ steht zum Diffusionskoeffizienten des Jodes im AgJ in naher Beziehung. Dieser gibt die Jodmenge an, welche durch eine AgJ-Schicht von 1 cm² und der Dickeneinheit an der das Konzentrationsgefälle 1 herrscht, in der Zeiteinheit diffundiert, während $2p$ proportional der J-Menge ist, die unter sonst gleichen Bedingungen, aber bei einem vom Werte 1 verschiedenen Konzentrationsgefälle durch 1 cm² der AgJ-Schicht diffundiert. Bezeichnet D den Diffusionskoeffizienten, a einen leicht zu berechnenden Reduktionsfaktor für die Umrechnung von AgJ-Dicken in Jod-Mengen und $\frac{dx}{dy}$ das Konzentrationsgefälle des Jods, so ist:

$$2p = aD\frac{dx}{dy}.$$

Würde man für zwei verschiedene Konzentrationsgefälle, zwei verschiedene bekannte Jod-Konzentrationen im Luftstrom die Werte $2p$ ermitteln, so könnte auch D berechnet werden.

Im Versuch 1) ist $2p$ merklich kleiner als im Versuch 2), obwohl bei beiden Versuchen die Oberflächen etwa gleich glatt waren. Auf der Oberfläche des Plättchens 1) befanden sich aber mehr Fremdkörper als auf der Oberfläche des Plättchens 2); diese erniedrigen, indem sie in die AgJ-Schicht übergehen, den Diffusionskoeffizienten von Jod im AgJ. Aus diesen Versuchen ist auf eine hohe Permeabilität des AgJ für Jod zu schließen, was sich auch

leicht auf anderem Wege bestätigen läßt. Bringt man einige Stückchen AgJ, gefällt und getrocknet oder geschmolzen, über Jodkristalle, so färben sich diese in einer Minute in der farblosen Luft bräunlich und nach 12 Stunden sind sie durch und durch braun gefärbt.

Man könnte der Meinung sein, daß die Abnahme der Verdickungsgeschwindigkeit auf eine Veränderung der AgJ-Schicht, also auf eine Alterungserscheinung, zurückzuführen ist, wodurch sich ihre Permeabilität ändern könnte. Diese Annahme kann experimentell leicht geprüft werden. Unterbricht man den Versuch, nachdem sich auf dem Plättchen eine Schicht von 150 $\mu\mu$ gebildet hat, und setzt ihn nach einer Stunde wieder fort, so verdickt sich die AgJ-Schicht innerhalb der Versuchsfehler in derselben Weise, als wenn die Unterbrechung nicht stattgefunden hätte. Auch ist es gleichgültig, ob das Plättchen, während der Unterbrechung im Licht oder im Dunkeln aufbewahrt wurde.

Da die Konstante $2p$ bei unveränderlichem Konzentrationsgefälle oder unveränderlicher Konzentration des Agens proportional dem Diffusionskoeffizienten ist, so sollte sie sich wie dieser mit der Temperatur ändern. Wenn also D proportional der absoluten Temperatur wächst, $D = c\,T$, so sollte dasselbe auch für $2p$ zutreffen. Erwärmt man die jodhaltige Luft und läßt sie auf ein Ag-Plättchen einwirken, so nimmt wegen der Ausdehnung der Luft die Jodkonzentration umgekehrt proportional der absoluten Temperatur ab, und dasselbe gilt für das Konzentrationsgefälle, $\frac{dx}{dy} = \frac{b}{T}$. In diesem Falle kompensieren sich der Temperatureinfluß auf den Diffusionskoeffizienten und das Konzentrationsgefälle, und $2p$ wird unabhängig von der Temperatur. Wenn also ein Einfluß der Temperatur auf die Verdickungsgeschwindigkeit der Anlaufschicht nicht festzustellen ist, so folgt hieraus, daß $D = c\,T$.

Um über den Einfluß der Temperatur auf den Vorgang des Anlaufens sich zu unterrichten, wurde ein konstanter Luftstrom durch die Lösung, enthaltend 0,002 Mol. J_2 + 0,014 Mol. KJ gesogen, der dann in eine Glasspirale von 2 m Länge trat, an deren Ende sich das Ag-Plättchen befand. Die Glasspirale wurde in ein Wasserbad gebracht, dessen Temperatur abwechselnd auf 15° und 75° gehalten wurde. Der zeitliche Verlauf des Erscheinens der Anlauffarben war in beiden Fällen im Wesentlichen derselbe. In fünf Versuchen war zu derselben Zeit vom Beginn des Versuchs die Anlauffarbe um höchstens zwei Stufen der Farbenskala ver-

schieden, es kam aber ebenso häufig vor, daß bei 15^{0} die Dicke der Anlaufschicht um höchstens 25 $\mu\mu$ die bei 70^{0} übertraf, als daß die umgekehrte Beziehung eintrat. Auf die Verdickungsgeschwindigkeit einer AgJ-Schicht hat also die Temperatur keinen merklichen Einfluß.

Die Beobachtung von Anlauffarben gestattet in jedem Moment die Dicke des entstandenen Reaktionsproduktes anzugeben und damit auch die Zahl der Atomschichten. Die erste Färbung, die auf einer Ag-Platte deutlich zu erkennen ist, ist ein helles Gelb entsprechend 63 . 10^{-6} mm. Die allerersten Farben, hell lavendelgrau bis grünlich, welche auftreten sollten, habe ich nie wahrnehmen können, ihre Wahrnehmung würde uns auch nicht zu viel dünneren Schichten, nämlich nur bis etwa 50 . 10^{-6} mm, führen. Eine AgJ-Schicht von 63 . 10^{-6} mm besteht aus etwa 90 Atomschichten oder 45 Atompaarschichten, wenn man in erster Annäherung den Gitterparameter des AgJ zu 7 . 10^{-7} mm annimmt. Von dieser Dicke bis zu etwa 900 Atomschichten reicht das Meßbereich der Methode der Farben dünner Schichten.

Anlauffarben entstehen nur dann, wenn bei der Einwirkung zusammenhängende Schichten entstehen. Sammelt sich das Reaktionsprodukt zu submikroskopischen Kristallen, so können Anlauffarben nicht entstehen. Ebenso können sie nicht entstehen, wenn die Dicke der zusammenhängenden Schicht nicht über etwa 50 Atomschichten im Reaktionsprodukt beträgt. Wenn wie bei der Einwirkung von Jod auf Cd eine Einwirkung überhaupt nicht zu beobachten ist, obwohl Cd und Jod unter Wasser sich schnell vereinigen, so kann der Grund hierfür entweder darin liegen, daß die gebildete CdJ_2-Schicht für J_2 undurchlässig ist, oder darin, daß sich überhaupt nicht eine CdJ_2-Schicht bildet, sondern die Cd-Atome der Oberfläche in ihrem Gitterverbande verbleiben, sich aber mit Jod-Atomen beladen. Die in diesem Falle gebildete Schicht ist jedenfalls nicht mit einer CdJ_2-Schicht identisch. Da CdJ_2-Kristalle im Jod-Dampf sich nicht färben, sich also in dieser Beziehung ganz anders als AgJ verhalten, so ist der Grund der Nichteinwirkung von Jod-Dampf auf Cd in der Bildung einer für Jod undurchlässigen Schicht zu suchen, während die Passivität des Fe wohl durch Bildung einer Zwitterschicht verursacht wird.

Die Atomistik läßt zwei verschiedene Arten der Diffusion im Raumgitter unterscheiden. Die Moleküle des eindringenden Stoffes können sich zwischen den Atomen des Gitters in dasselbe hineinzwängen, oder, wenn zwei Atomarten das Gitter besetzen, so könnte beim Aufprall der einen Atomart auf das Gitter der Stoß

im Gitter sich fortpflanzen bis auf einer nur mit dieser Atomart besetzten Gittergeraden das letzte Atom frei wird. Diese Stoßdiffusion sollte in sehr viel höherem Maße als die gewöhnliche Diffusion von der Richtung im Gitter abhängen.

Die Geschwindigkeit der Entstehung von Anlauffarben wird also nur durch den Diffusionskoeffizienten des Agens in der Schicht des sich bildenden Reaktionsproduktes bestimmt, die Geschwindigkeit der Bildung des Reaktionsproduktes selbst macht sich nicht geltend, weil sie sehr viel größer als die Diffusionsgeschwindigkeit ist.

Obwohl auf die Geschwindigkeit der Entstehung der Anlauffarben des Ag in Jod-Dampf das Licht keinen merklichen Einfluß ausübt, und auch für die Anlauffarbe nach ihrer Entstehung im allgemeinen dasselbe gilt, so ist doch auf die Farbe ganz bestimmter unter bestimmten Bedingungen entstandenen Anlauffarben ein starker Einfluß des Lichtes, der allerdings ganz unaufgeklärt geblieben ist, festzustellen. Läßt man die Oberseite des Ag-Plättchens bis zum Rot IV. Ordnung anlaufen, so zeigt der mittlere Teil der unteren Oberfläche noch keine Farbe, weil die Strömungsgeschwindigkeit der Luft hier recht gering ist. Dieser farblose Fleck ist von einem braungelben Ringe, dem ein himmelblauer folgt, umgeben, diesem folgen dann in weiterer Entfernung 3 Ringe mit den Farben III., IV. und V. Ordnung. Bei stärkerer Beleuchtung wird der gelbbraune Ring tief braunschwarz und der himmelblaue tief dunkelblau, beide verengen sich und schließlich ist die Mitte schwach rötlich gefärbt, umgeben von einem grünen Ringe, während die drei folgenden Farbenringe weder ihre Lage noch ihre Farbe merklich geändert haben. Die Geschwindigkeit dieses Vorganges wächst mit der Intensität des Lichtes, im Sonnenlicht ist der Endzustand in einer Stunde erreicht, während er im zerstreuten Tageslicht 10 bis 20 Stunden zu seiner Vollendung braucht. Im jodhaltigen Luftstrom wird der grüne Ring zuerst gelb, dann tiefblau und gelb, dann erst beginnt die rötliche Mitte ihre Farbe in gelb, rot, violett und grün zu ändern. Es treten abnorme Färbungen und in abnormer Reihenfolge auf, sodaß man die merkwürdige Erscheinung nicht enträtseln kann.

Die Geschwindigkeit des Anlaufens von Stahl in ihrer Abhängigkeit von der Zeit und der Temperatur.

Diesen Vorgang hat L. Löwenherz [1]) eingehender untersucht, indem er flache, polierte Stahlstäbe auf Metallbädern bekannter Temperatur schwimmen ließ und die Zeiten bestimmter Anlauffarben feststellte. Leider können die isothermen Versuche bei 177, 191 und 209° nur bis zu den Farben zweiter Ordnung ausgedehnt werden, weil die Farben höherer Ordnung schon sehr blaß sind, und weil in diesem Temperaturintervall sich die Farben erst nach Tagen oder Stunden ändern. Rechnet man für diese Versuche die Konstante $2p$ aus, so ergeben sich Werte, welche innerhalb der Fehlergrenzen als unveränderlich angesehen werden können. Allerdings können aus den beobachteten Anlauffarben nicht die Dicken der Anlaufschicht ermittelt werden, sondern nur Dicken der Luftschichten, die den Farben entsprechen. Da aber diese Dicken y den Dicken βy_1 der Anlaufschichten proportional sind, so ergibt sich aus:

$$y^2 = 2pt, \quad y_1^2 = \frac{2p}{\beta^2} t.$$

Die Konstante $2p$ wäre also, wenn sie sich auf die Dicke der Anlaufschicht selbst beziehen sollte, noch durch den unveränderlichen Wert β^2 zu dividieren. Wollte man die Temperaturabhängigkeit von $2p$ genau ermitteln, so hätte man $2p$ für eine Reihe von Temperaturen aus isothermen Versuchen zu ermitteln. Derartige Versuche sind nun nicht angestellt, aber durch einen günstigen Umstand, ist noch eine andere Reihe nicht isothermer Versuche berechenbar geworden. Löwenherz ließ nämlich in einem auf 249° erhitzten Metallbade die Stahlplatten so schwimmen, daß er ihre polierte Oberfläche beobachten konnte und erhitzte nun das Bad. Wenn $2p$ von der Temperatur unabhängig ist, so kann sein Wert ohne Bedenken auch aus diesen Beobachtungen ermittelt werden. In der folgenden Tabelle sind die Zeiten t, gerechnet vom Momente des Eintauchens der Stahlplatte in das Bad, dessen Temperatur zur Zeit t und die der beobachteten Farbe entsprechende Dicke d der Luftschicht nebst den berechneten Werten von $2p$, zusammengestellt für den Versuch an weicherem Stahl.)

1) Zeitschrift für Instrumentenk. 1889, S. 316.

Zeit in Minuten	20	21	25	28	30	32	35	39	45	50	56
Temperatur	249	258	287	315	325	330	340	353	362	379	392
$d \cdot 10^6$ mm	140	164	235	257	272	282	300	387	409	435	490
$2p \cdot 10^{11}$	98	128	221	236	247	248	257	384	372	378	437

Aus den isothermen Versuchen ergibt sich $2p$ bei 177° zu $1{,}35 \cdot 10^{-11}$ und bei 204° zu $1{,}67 \cdot 10^{-11}$, es wächst also zwischen 204° und 249° $2p$ auf das 60fache seines Wertes, während unterhalb 204° die Änderung von $2p$ sehr viel geringer ist. Zwischen 290° und 340° ändert sich $2p$ nur sehr wenig mit der Temperatur, wächst bei 350° nochmals sprungweise, wenn auch längst nicht in dem Maße wie zwischen 204 und 249° und scheint sich dann wiederum wenig mit der Temperatur zu ändern. Für gehärteten Stahl ergab sich, daß $2p$ zwischen 325 und 357° unabhängig von der Temperatur ist; in diesem Intervall ergeben sich für $2p \cdot 10^{-11}$ die Werte 184, 188, 190 und 194.

Auch hier finden sich wie bei der Einwirkung von Jod auf Ag Temperaturgebiete, in denen die Verdickungsgeschwindigkeit der Anlaufschicht unabhängig von der Temperatur ist. Da die Konzentration des Sauerstoffs in der erhitzten Luft umgekehrt proportional der absoluten Temperatur ist, so muß der Diffusionskoeffizient des Sauerstoffs in der Anlaufschicht in diesen Temperaturgebieten proportional der absoluten Temperatur sein. Außerdem ändert nach diesen Versuchen die Anlaufschicht ihren Diffusionskoeffizienten zweimal fast diskontinuirlich, was auf eine Änderung der Anlaufschicht selbst hinweist. Ob diese in dem Übergange eines niederen Eisenoxydes in ein höheres oder in der Bildung einer anderen Form desselben Oxydes zu suchen ist, muß dahingestellt bleiben. Jene starken Änderungen des Diffusionskoeffizienten treten allerdings nicht bei denselben Temperaturen ein, doch sind Verzögerungen bei den Umwandlungen anisotroper Phasen bekanntlich nicht selten und gerade bei kleinen Mengen der Phase zu erwarten.

Die Konstitution des Cholesterins.

Von

A. Windaus.

Vorgelegt in der Sitzung vom 30. Mai 1919.

Um die Aufklärung der Konstitution des Cholesterins haben sich schon seit Jahrzehnten die Chemiker vielfach bemüht, indessen haben erst die mit Beginn dieses Jahrhunderts von verschiedenen Seiten in Angriff genommenen zielbewußten und umfangreichen Untersuchungen die ersten tiefer gehenden Erfolge erzielt. Doch ist auch heute der Aufbau des Kohlenstoffgerüstes im Cholesterin noch nicht klar gelegt. Dies liegt in erster Linie an der großen Zahl (27) von Kohlenstoffatomen, die im Molekül des Cholesterins unmittelbar miteinander verbunden sind, ferner daran, daß Stoffe mit komplizierten hydrierten Ringsystemen — und dazu gehören Cholesterin und wahrscheinlich viele andere Stoffe des Tier- und Pflanzenreichs — bisher synthetisch nicht zugänglich sind und daher systematisch kaum studiert sind, endlich daran, daß bei energischen Abbauversuchen am Cholesterin meist eine ganze Anzahl von Umwandlungsprodukten entstehen, die sich gegenseitig an der Krystallisation hindern und deren sichere Trennung und Reindarstellung oft nicht gelingt. So kommt es, daß trotz zäher und mannigfaltiger experimenteller Arbeit die Erkenntnis hier nur langsam fortschreitet und daß auch mancher Irrtum in der Deutung des experimentellen Materials vorgekommen ist.

Formel.

Schon die Feststellung der richtigen Formel für das Cholesterin hat große Schwierigkeiten bereitet, weil bei einem so hochmolekularen Stoff die Berechnung der Formel aus den Analysenzahlen

vieldeutige Ergebnisse liefert. Die von Berthelot[1]) aufgestellte Formel $C_{26}H_{44}O$ verlangt 83.80 % C, 11.91 % H; die jetzt angenommene $C_{27}H_{46}O$ 83.86 % C, 11.99 % H. Größere Differenzen in den Analysenzahlen finden sich bei den halogenhaltigen Derivaten des Cholesterins, und durch die Analyse solcher hat Reinitzer[2]) als erster die Formel $C_{27}H_{46}O$ sehr wahrscheinlich gemacht. Viele Mikroanalysen von Pregl und Lieb haben die Richtigkeit dieser Formel bewiesen. Mauthner[3]) hat allerdings noch in seinen letzten Arbeiten die Formel $C_{27}H_{44}O$ bevorzugt und auch alle Derivate von dieser Formel abgeleitet. Dasselbe habe ich in meinen ersten Arbeiten aus den Jahren 1903 bis 1906 getan. Hierdurch ist es gekommen, daß eine gewisse Unsicherheit über die Zahl der Wasserstoffatome in Cholesterinderivaten besteht und in vielen Fällen die in der Literatur angegebene Formel um zwei Wasserstoffatome vermehrt werden muß.

Hydroxylgruppe.

In der Formel $C_{27}H_{46}O$ befindet sich ein Sauerstoffatom; die ersten chemischen Arbeiten über Cholesterin befassen sich damit die Funktion dieses Sauerstoffatoms zu bestimmen. Schon Berthelot[4]) hat aus dem Cholesterin durch Erhitzen mit organischen Säuren Ester erhalten und so bewiesen, daß Cholesterin ein Alkohol ist. Auch zahlreiche andere Umsetzungen der Hydroxylgruppe sind studiert worden, aber erst 1904 ist es Diels und Abderhalden[2]) gelungen das Cholesterin durch Erhitzen mit Kupferoxyd auf 280—300° zu dem entsprechenden Keton, dem Cholestenon, zu oxydieren: Cholesterin ist ein sekundärer Alkohol.

Doppelbindung.

Über den Sättigungszustand des Cholesterinmoleküls haben zuerst Wislicenus und Moldenhauer[6]) Versuche angestellt und erkannt, daß Cholesterin eine Doppelbindung enthält, da es glatt ein Molekül Brom zu addieren vermag. Auch durch Anlagerung von Ozon hat man versucht über die Zahl der Doppel-

1) Ann. chim. phys. **56**, 51 (1856).
2) Monatsh. f. Chem. **9**, 421 (1888).
3) Monatsh. f. Chem. **30**, 647 (1909).
4) loc. cit.
5) Berichte D. chem. Ges. **37**, 3099 (1904).
6) Liebigs Annalen **146**, 175 (1868).

bindungen im Cholesterin Aufschluß zu gewinnen; diese Methode ist aber recht unsicher[1]). Die Anlagerung von Wasserstoff an Cholesterin ist häufig vergeblich versucht worden; viele Reduktionsmittel wirken garnicht ein; die Behandlung mit Natrium und Amylalkohol liefert ein Amylkondensationsprodukt[2]); erst die Ausarbeitung der katalytischen Hydrierungsmethoden hat es ermöglicht an Cholesterin glatt 1 Mol. Wasserstoff zu addieren. Hierbei entsteht das Dihydrocholesterin oder β-Cholestanol, ein gesättigter sekundärer Alkohol von der Formel $C_{27}H_{48}O$, der sich leicht in den gesättigten Stammkohlenwasserstoff der Cholesterinreihe, das Cholestan $C_{27}H_{48}$ umwandeln läßt[4]).

Ringsysteme im Cholesterin.

Das Cholestan $C_{27}H_{48}$ enthält 8 Wasserstoffatome weniger als das gesättigte Paraffin mit 27 Kohlenstoffatomen $C_{27}H_{56}$; dieses Minus von 8 Wasserstoffatomen muß bedingt sein durch das Vorhandensein von 4 hydrierten Ringen im Molekül des Cholestans.

Cyklische Bindung der Alkoholgruppe.

Einer dieser Ringe läßt sich sehr leicht beim oxydativen Abbau des Cholesterins nachweisen. Im Jahre 1903 und 1904 haben Diels und Abderhalden[5]) gefunden, daß, wenn man Cholesterin mit unterbromigsaurem Kalium oxydiert, eine Dicarbonsäure $C_{27}H_{44}O_4$ gebildet wird; es findet also eine Aufspaltung an der Hydroxylgruppe statt, ohne daß das Molekül in kleinere Bruchstücke zerfällt. Der Vorgang ist durch die folgende Formel wiederzugeben

$$\overbrace{\underset{CH_2}{|}\text{—}\underset{CHOH}{|}}^{C_{25}H_{42}} + 4O = \overbrace{\underset{COOH}{|}\quad\underset{COOH}{|}}^{C_{25}H_{42}} + H_2O$$

und beweist, daß im Cholesterin eine cyklisch gebundene $-\overset{H_2}{C}\text{—}C{<}^{H}_{OH}$ -Gruppe vorhanden ist, aus der bei der Oxydation zwei Carboxyle hervorgehen.

1) Fürth und Felsenreich, Bioch. Ztschr. **69**, 416 (1915).

2) Windaus und Uibrig, Berichte D. chem. Ges. **46**, 2487 (1913).

3) Willstätter und Mayer, Berichte D. chem. Ges. **41**, 2199 (1908). S. auch Diels und Abderhalden, Berichte D. chem. Ges. **39**, 889 (1906).

4) Diels und Linn, Berichte D. chem. Ges. **41**, 548 (1908) und Mauthner, Monatsh. f. Chemie **30**, 638 (1909).

5) Berichte D. chem. Ges. **36**, 3179 (1903) und **37**, 3092 (1904).

Auch an vielen Cholesterinderivaten, die die sekundäre Alkoholgruppe noch enthalten, ist dieselbe Aufspaltung vorgenommen worden. Sehr glatt läßt sich z. B. Dihydrocholesterin $C_{27}H_{48}O$ über das entsprechende gesättigte Keton, das β-Cholestanon $C_{27}H_{46}O$ zu der Dicarbonsäure $C_{27}H_{46}O_4$ oxydieren[1]).

Cyklische ungesättigte Gruppe.

Daß auch die ungesättigte Gruppe des Cholesterins in einem zweiten Ringe steht, ist zuerst durch Untersuchungen von Windaus und Stein[2]) ermittelt worden. Als Ausgangsmaterial für diese Versuche haben die ungesättigten Nitroverbindungen gedient, die man aus Cholesterin und seinen Derivaten mittels rauchender Salpetersäure erhält. Aus Cholesterin entsteht so, indem gleichzeitig die Hydroxylgruppe durch die Salpetersäure verestert wird, ein Nitrocholesterylnitrat $C_{27}H_{44}(NO_2)(ONO_2)$; aus Cholesterylacetat $C_{27}H_{45}(OCOCH_3)$ ein Nitrocholesterylacetat $C_{27}H_{44}(NO_2)(OCOCH_3)$[3]); aus Cholesterylchlorid $C_{27}H_{45}Cl$ ein Nitrocholesterylchlorid[4]) $C_{27}H_{44}(NO_2)(Cl)$; endlich aus dem Kohlenwasserstoff Cholesten[5]) $C_{27}H_{46}$ (dargestellt aus Cholesterylchlorid mit Natrium und Amylalkohol) das Nitrocholesten $C_{27}H_{45}(NO_2)$[6]). Bei dieser Reaktion wird vermutlich ein an einem ungesättigten Kohlenstoffatom haftendes Wasserstoffatom durch die Nitrogruppe ersetzt. Analog verlaufende Reaktionen sind vielfach studiert worden, so z. B. am Methylinden[7]):

C—CH$_3$ ‖ CH, CH$_2$ → C—CH$_3$ ‖ C—NO$_2$, CH$_2$.

Diese Nitroverbindungen geben nun bei der Reduktion mit Zinkstaub und Essigsäure gesättigte Ketone, das Nitro-inden liefert β-Hydrindon[8]), das Nitrocholesterylnitrat liefert, indem gleichzeitig der Salpetersäureester verseift wird, einen Ketoalkohol, das Cholestanon-ol[9]) $C_{25}H_{44}(CHOH)(CO)$; das Nitro-

1) Windaus und Uibrig Berichte D. chem. Ges. **47**, 2387 (1914).
2) Berichte D. chem. Ges. **37**, 3699 (1904).
3) Mauthner und Suida Monatsh. f. Chemie **24**, 652 (1904).
4) Ebd. **15**, 104 (1894). 5) Ebd. 86 (1894). 6) Ebd. 109 (1894).
7) Wallach, Liebigs Annalen, **336**, 5 (1904).
8) Ebd. 2 (1904)
9) Windaus, Berichte D. chem. Ges. **36**, 3753 (1903).

cholesten liefert ein Keton $C_{26}H_{46}$ (CO), das den Namen Heterocholestanon erhalten hat. Das Cholestanon-ol, das also an Stelle des —CH : C-Restes die Gruppierung (—CO—CH—) enthält, hat sich zu mannigfachen Umsetzungen als brauchbar erwiesen, So läßt sich die Hydroxylgruppe mittels Phosphorpentachlorid durch Chlor ersetzen und dann das gebildete Chlorcholestanon[1]) $C_{26}H_{45}$ (CO) Cl mit Salpetersäure zu einer Chlordicarbonsäure $C_{27}H_{45}ClO_4$ [2]) oxydieren:

$$\overbrace{\underset{CH_2}{|}—\underset{CH\,OH}{|}\quad\underset{CO}{|}—\underset{CH_2}{|}}^{C_{23}H_{40}} \rightarrow \overbrace{\underset{CH_2}{|}—\underset{CH\,Cl}{|}\quad\underset{CO}{|}—\underset{CH_2}{|}}^{C_{23}H_{40}} \rightarrow$$

$$\rightarrow \overbrace{\underset{CH_2}{|}—\underset{CH\,Cl}{|}\quad\underset{COOH}{|}\quad\underset{COOH.}{|}}^{C_{23}H_{40}}$$

Die Entstehung dieser Dicarbonsäure, die zum Chlorcholestanon in demselben Verhältnis steht wie Kamphersäure zu Kampher, beweist, daß die CO-Gruppe im Chlorcholestanon in einem hydrierten Ringe und neben einer CH_2-Gruppe steht. Hieraus folgt, daß auch die Doppelbindung, aus der ja die Ketogruppe hervorgegangen ist, ringförmig gebunden ist.

Daß dieser neue Ring (2) verschieden ist von dem Ring (1), der die Hydroxylgruppe trägt, läßt sich leicht zeigen. Wird nämlich in der Chlordicarbonsäure $C_{27}H_{45}ClO_4$ an Stelle des Chlors das ursprünglich vorhandene Hydroxyl wieder eingeführt, erhält man eine Oxysäure $C_{27}H_{46}O_4$, die sich zu einer Ketosäure $C_{27}H_{44}O_4$ und weiter unter Sprengung von Ring (1) zu einer Tetracarbonsäure $C_{27}H_{44}O_8$ oxydieren läßt:

$$\underset{\text{Oxysäure}}{\overbrace{\underset{CH_2}{|}\overset{1)}{—}\underset{CHOH}{|}\quad\underset{COOH}{|}\quad\underset{COOH}{|}}^{C_{23}H_{40}}} \rightarrow \underset{\text{Ketosäure}}{\overbrace{\underset{CH_2}{|}\overset{1)}{—}\underset{CO}{|}\quad\underset{COOH}{|}\quad\underset{COOH}{|}}^{C_{23}H_{40}}} \rightarrow$$

$$\rightarrow \underset{\text{Tetracarbonsäure}}{\overbrace{\underset{COOH}{|}\quad\underset{COOH}{|}\quad\underset{COOH}{|}\quad\underset{COOH.}{|}}^{C_{23}H_{40}}}$$

Anmerkung: Aus dem eben mitgeteilten Tatsachenmaterial ergibt sich also, daß Hydroxylgruppe und Doppelbindung des

1) Windaus und Stein, Berichte D. chem. Ges. 37, 3702 (1904).
2) Ebd. 3704 (1904).
3) Ebd. 3704 (1904).

Cholesterins in zwei verschiedenen hydrierten Ringen stehen. An der Richtigkeit dieser Schlußfolgerung habe ich allerdings zeitweise gezweifelt und die Vermutung ausgesprochen, daß der Ring (2), der die aus der Doppelbindung hervorgegangene Ketogruppe trägt, ein Kunstprodukt sei; ich glaubte nämlich aus den Analysen von Mauthner und Suida schließen zu müssen, daß bei der Bildung der Nitroverbindungen zwei Wasserstoffatome aus dem Molekül des Cholesterins abgespalten und Dehydroderivate gebildet würden, und verglich die Bildung dieser Dehydroprodukte mit dem Übergang von Akroleinanilin in Chinolin und nahm auf Grund einiger Oxydationsversuche an, daß eine ursprünglich vorhandene entständige Vinylgruppe unter Abspaltung von zwei Wasserstoffatomen in eine cyklisch gebundene —CH : CH-Gruppe verwandelt worden sei. Die folgenden Formeln sollen diese (inzwischen als falsch erkannte) Annahme veranschaulichen.

$$\begin{array}{c} C_{23}H_{39}—CH:CH_2 \\ /\backslash \\ CH_2—CHOH \end{array} \rightarrow \begin{array}{c} \quad\quad /CH \\ C_{23}H_{38}\langle\ \ \| \\ /\backslash \quad \backslash CH \\ CH_2—CHOH \end{array} \rightarrow \begin{array}{c} \quad\quad /C—NO_2 \\ C_{23}H_{38}\langle\ \ \| \\ /\backslash \quad \backslash CH \\ CH_2—CHOH \end{array}$$

$$\rightarrow \begin{array}{c} \quad\quad /CO \\ C_{23}H_{38}\langle\ \ | \\ /\backslash \quad \backslash CH_2. \\ CH_2—CHOH \end{array}$$

Später habe ich aber völlig einwandfrei feststellen können [3]), daß diese Annahme einer künstlichen Ringbildung nicht richtig ist; daß vielmehr in den Nitroderivaten und ihren Umwandlungsprodukten noch dieselbe Anordnung der Kohlenatoffatome vorhanden ist wie im Cholesterin; dies folgt daraus, daß sich das Cholestanon-ol und das Heterocholestanon zu demselben gesättigten Kohlenwasserstoff, dem Cholestan reduzieren lassen, der auch aus dem Cholesterin entsteht. Die beiden in Cholestanon-ol nachgewiesenen alicyklischen Systeme sind also auch im Cholesterin vorhanden.

Ueber die gegenseitige Stellung von Hydroxylgruppe und Doppelbindung.

Durch die bisherigen Versuche sind also im Cholesterin eine sekundäre Alkoholgruppe und eine ungesättigte Gruppe CH : C

---

1) Berichte D. chem. Ges. **39**, 2262 (1906).
2) Archiv d. Pharmacie **246**, 134 (1908).
3) Berichte D. chem. Ges. **50**, 134 (1917).

bezw. CH : CH nachgewiesen, die in zwei verschiedenen hydrierten Ringen stehen. Einige weitere Abbauversuche gestatten auch gewisse Schlüsse über die gegenseitige Stellung von Alkoholgruppe und Doppelbindung zu ziehen. Wird nämlich das Cholestenon, das Keton des Cholesterins von der Formel

$$\overbrace{\overset{|}{C}H_2-\overset{|}{C}O \quad \overset{|}{C}H:\overset{|}{C}H}^{C_{23}H_{40}}$$

mit wässeriger Kaliumpermanganatlösung vorsichtig oxydiert[1]), bleibt die Ketogruppe unverändert, während die ungesättigte Gruppe angegriffen wird. Statt der erwarteten Ketodicarbonsäure

$$\overbrace{\overset{|}{C}H_2-\overset{|}{C}O \quad \overset{|}{C}OOH \quad \overset{|}{C}OOH}^{C_{23}H_{40}}$$

erhält man die Bruchstücke, Kohlendioxyd und eine Ketomonocarbonsäure $C_{26}H_{44}O_3$; die zunächst gebildete Ketodicarbonsäure $C_{27}H_{44}O_5$ ist augenscheinlich sehr unbeständig und spaltet leicht Kohlendioxyd ab, sie enthält daher wahrscheinlich die eine Carboxylgruppe in β-Stellung zur Ketogruppe. Versucht man den ganzen Reaktionsverlauf durch ein Formelbild wiederzugeben, kommt man nur zu einer einleuchtenden Deutung, wenn man annimmt, daß sich die Doppelbindung im Cholestenon in β, γ-Stellung zur Ketogruppe befindet:

$$\overbrace{\overset{|}{C}H_2 \quad \overset{|}{C}H \quad \overset{|}{C}H}^{C_{22}H_{39}} \quad (\text{CH}_2\text{ und CH über CO verbunden; CH und CH über CH verbunden}) \rightarrow \overbrace{\overset{|}{C}H_2 \quad \overset{|}{C}H \quad \overset{|}{C}OOH}^{C_{22}H_{39}} \quad (\text{CH}_2\text{ und CH über CO verbunden; CH trägt COOH})$$

Cholestenon

$$\rightarrow \overbrace{\overset{|}{C}H_2 \quad \overset{|}{C}H_2 \quad \overset{|}{C}OOH}^{C_{22}H_{39}} + CO_2. \quad (\text{CH}_2\text{ und CH}_2\text{ über CO verbunden})$$

Säure $C_{26}H_{44}O_3$

Anmerkung: Auch hier ist anfangs der Reaktionsverlauf nicht richtig erkannt worden: ich habe zuerst der Ketomonocarbonsäure $C_{26}H_{44}O_3$ eine um zwei Wasserstoffatome ärmere Formel zuerteilt

1) Windaus, Berichte D. chem. Ges. 39, 2010 (1896) und 50, 133 (1917),

und den Vorgang unter Annahme einer endständigen Vinylgruppe im Cholestenon folgendermaßen gedeutet:

$$\underbrace{\begin{matrix} C_{22}H_{39}-CH:CH_2 \\ | \quad\quad | \\ CH_2-CO \end{matrix}} + 5O = \begin{matrix} C_{22}H_{39}-COOH \\ | \quad\quad | \\ CH_2-CO \end{matrix} + H_2O + CO_2.$$

Erst der bewährten analytischen Kunst des Herrn Professor Pregl und des Herrn Dr. H. Lieb Graz ist es geglückt die Formel $C_{26}H_{44}O_8$ mit Sicherheit zu beweisen und dadurch die richtige Erklärung für den Reaktionsverlauf zu ermöglichen.

Die β, γ-Stellung der Doppelbindung habe ich noch auf einem anderen Weg nachzuweisen gesucht. Wie oben erwähnt, läßt sich im Cholesterin

$$\begin{matrix} & & C_{22}H_{39} & & \\ | & & | & & | \\ CH_2 & & CH & & CH \\ & \diagdown\;\diagup & & \diagdown\;/\!\!/ & \\ & CHOH & & CH & \end{matrix}$$

die Gruppe CH : CH durch $COCH_2$ ersetzen; das dabei gebildete Cholestanon-ol liefert mit Chromsäure oxydiert ein Diketon das Cholestandion[1]), das also ein β- oder γ-Diketon

$$\begin{matrix} & & C_{22}H_{39} & & \\ | & & | & & | \\ CH_2 & & CH & & CH_2 \\ & \diagdown\;\diagup & & \diagdown\;\diagup & \\ & CO & & CO & \end{matrix} \quad \text{oder} \quad \begin{matrix} & & C_{22}H_{39} & & \\ | & & | & & | \\ CH_2 & & CH & & CO \\ & \diagdown\;\diagup & & \diagdown\;\diagup & \\ & CO & & CH_2 & \end{matrix}$$

sein muß, wenn Cholesterin ein β, γ-ungesättigter Alkohol ist. Für β-(1,3) und γ-(1,4) Diketone ist es charakteristisch, daß sie mit 1 Mol. Hydrazin unter Austritt von 2 Mol. Wasser reagieren und dabei Pyrazol- bezw. Pyridazinderivate liefern. Ganz so verhält sich nun das Cholestandion, das beim Erhitzen mit Hydrazin ein Kondensationsprodukt $C_{27}H_{44}N_2$ [2]) liefert:

$$C_{27}H_{44}O_2 + N_2H_4 = C_{27}H_{44}N_2 + 2H_2O.$$

Die weitere Untersuchung hat dann ergeben, daß Cholestandion sicher kein 1,3 Diketon ist; es ist unlöslich in Alkalien, es liefert mit Hydroxylamin ein normales Dioxim und kein Isoxazolderivat

---

1) Windaus, Berichte D. chem. Ges. **36**, 3755 (1903).
2) Ebd. **39**, 2256 (1906).

und gibt bei der Oxydation mit Chromsäure, wahrscheinlich unter Aufspaltung des Ringes (1), eine Ketodicarbonsäure $C_{27}H_{44}O_5$ [1]), die keine Neigung besitzt Kohlendioxyd abzuspalten:

```
      C22 H39                        C22 H39
 _____________________        ________________________
 |        |         |          |          |        |
CH2      CH        CO  ->    COOH        CH       CO;
   \    /  \      /                     /  \     /
     CO      CH2                    COOH    CH2
```

ein 1,3 Diketon würde aber beim Abbau eine Ketodicarbonsäure liefern, deren eine Carboxylgruppe in β-Stellung zur Ketogruppe stehen und vermutlich leicht abspaltbar sein müßte [2]).

Endlich soll noch eine andere Reaktionsfolge angeführt werden, die ebenfalls für die 1,4 Stellung der Carbonylgruppen im Cholestandion spricht. Wird Cholesterin (I) mit Benzopersäure in Chloroformlösung oxydiert [3]), so entsteht ein Oxyd (II), das durch Wasser zu einem Cholestantriol (III) aufgespalten wird:

```
      C22 H39                          C22 H39
 _____________________         __________________________
 |        |         |           |         |    O     |
CH2      CH        CH   ->    CH2        CH    |  \ CH
   \    /  \      //             \      /  \   |   /
    CHOH     CH                    CHOH      CH
      I.                              II.

                  C32 H29
            ____________________
            |        |        |
      ->   CH2      CH       CHOH.
              \    /  \     /
               CHOH    CHOH
                   III.
```

Dieses Cholestantriol wird durch Chromsäure zu einem Diketoalkohol, dem Cholestandion-ol (IV) oxydiert, das sehr leicht Wasser abspaltet und ein ungesättigtes Diketon, das Cholestendion oder Oxycholestenon (V) liefert. Dieses Oxycholestenon, dessen eine Ketogruppe in der Enolform reagiert und das auch direkt aus Cholesterin und Chromsäure entsteht [4]), wird durch Zinkstaub und Essigsäure ganz glatt zum Cholestandion (VI) reduziert [5]):

---

1) Windaus, Berichte D. chem. Ges. **36**, 3756 (1903).

2) Siehe indessen Ossian Aschan, Liebigs Ann. **410**, 243 (1915).

3) Westphalen, Berichte D. chem. Ges. **48**, 1064 (1915); s. auch Pickard und Yates, Journ. Chem. Soc. **93**, 1683 (1908).

4) Mauthner und Suide, Monatsh. f. Chemie **17**, 584 (1896).

5) Windaus, Berichte D. chem. Ges. **39**, 2252 (1906).

$C_{22}H_{39}$ — CH_2 CH CO / CO $CHOH$ (IV.) → $C_{22}H_{39}$ — CH_2 C CO / CO CH (V.) → $C_{22}H_{39}$ — CH_2 CH CO. / CO CH_2 (VI.)

Das Cholestendion (Oxycholestenon) addiert bei vorsichtiger Behandlung mit Brom 1 Mol. Brom und gibt ein schön krystallisiertes Dibromid $C_{27}H_{42}O_2Br_2$; das Cholestandion dagegen wird durch Brom substituiert und liefert ein Dibromid, das mit dem Produkt aus Oxycholestenon identisch ist; in dem Substitutionsprodukt werden sich die Bromatome in der α-Stellung zu den Carbonylgruppen befinden, in dem Additionsprodukt werden sie benachbart stehen. Diesen beiden Forderungen kann nur eine Formulierung genügen, bei der sich die beiden Ketogruppen in 1,4 Stellung und die Doppelbindung des Oxycholestenons in der Mitte zwischen den beiden Carbonylgruppen befindet[1]).

Ueber die Gliederzahl der Ringe im Cholesterin.

Während das Vorhandensein zweier aneinandergelagerter Ringe im Cholesterin durch die eben mitgeteilten Untersuchungen sicher gestellt ist, kann über die Gliederzahl dieser Ringe zunächst nichts ausgesagt werden. Ein Herausschälen des Ringsystems aus dem großen Cholesterinmolekül hat sich nicht ermöglichen lassen. Windaus und Dalmer[2]) haben darum einen indirekten Weg für die Untersuchung eingeschlagen: Es gelingt leicht im Cholesterin

$C_{22}H_{39}$ — CH_2 1) CH 2) CH / $CHOH$ CH

den Ring (2) durch Hydrierung an der Doppelbindung vor der Oxydation zu schützen und dann Ring (1) aufzuspalten; ebenso kann auch Ring (1) nach Ersatz der Hydroxylgruppe durch Wasserstoff vor der Oxydation geschützt und dann Ring (2) aufgespalten werden. In beiden Fällen entstehen gut charakterisierbare Dicarbonsäuren, in denen sich die gegenseitige Stellung der beiden Carboxylgruppen mit großer Wahrscheinlichkeit ermitteln läßt, und hieraus kann dann ein Schluß über die Zahl der Ringglieder in den ursprünglich vorhandenen Ringen gezogen werden. Um die

1) Windaus, Berichte D. chem. Ges. 39, 2254 (1906).
2) Berichte D. chem. Ges. 52, 162 (1919).

Stellung der Carboxylgruppe festzustellen, haben Windaus und Dalmer die Methode von Blanc[1]) verwendet, die auf dem verschiedenen Verhalten gesättigter Dicarbonsäuren gegenüber Essigsäureanhydrid beruht. Werden 1,4 oder 1,5 Dicarbonsäuren mit Essigsäureanhydrid auf 215—240° erhitzt, so geben sie die entsprechenden inneren Anhydride; 1,6 oder 1,7 Dicarbonsäuren liefern dagegen bei derselben Behandlung penta- oder hexacyklische Ketone. So erhält man aus Pimelinsäure Cyklohexanon, aus Adipinsäure Cyklopentanon. Die Reaktion führt also in diesem Falle zu denselben Produkten wie die trockne Destillation der Calcium- oder Bleisalze. 1,8 Dicarbonsäuren liefern nur wenig heptacyklisches Keton. Die Methode von Blanc ist bisher systematisch nur bei aliphatischen Dicarbonsäuren geprüft worden; es erhebt sich daher die Frage, ob die Methode auch in der hydroaromatischen Reihe zuverlässige Ergebnisse liefern wird. Dies darf mit großer Wahrscheinlichkeit angenommen werden, da ja auch bei bi- und polycyklischen Verbindungen die Bildung von Fünf- und Sechsringen begünstigt erscheint. Windaus und Dalmer haben daher geglaubt auch in der Cholesterinreihe die Blanc'sche Methode als diagnostisches Hilfsmittel verwenden zu können. Bei der Prüfung des Ringes 1 wurde zunächst Dihydrocholesterin

```
          C22 H39
  ________________________
  |         |          |
 CH2        CH         CH2
    \     /    \     /
     CHOH        CH2
```

bereitet und mittels Chromsäure zu der entsprechenden Dicarbonsäure

```
           C22 H39
  __________________________
  |           |          |
 COOH         CH         CH2
            /    \     /
        COOH       CH2
```

aufgespalten. Beim Erhitzen dieser Säure mit Essigsäureanhydrid und nachfolgender Destillation im Vakuum entsteht in fast quantitativer Ausbeute ein Keton $C_{26}H_{44}O$ oder

```
          C22 H39
  ______________________
  |        |          |
  |        |         CH2,
  |        |        /
 CO ——— CH ——— CH2
```

1) Compt. rend. 144, 1356 (1907).

das sehr schön krystallisiert. Die Säure verhält sich also wie eine Dicarbonsäure mit 1,6 oder 1,7 Stellung der Carboxyle. Das Keton läßt sich beim Erwärmen mit rauchender Salpetersäure wiederum glatt oxydieren und gibt dabei eine Dicarbonsäure $C_{26}H_{44}O_4$, deren Bildung beweist, daß sich in dem Keton eine Methylengruppe neben der Carbonylgruppe befindet:

$$\overbrace{\underset{\underset{CO}{\diagdown\;\diagup}}{CH_2\quad CH}\quad \underset{\underset{CH_2}{\diagdown\;\diagup}}{\quad CH_2}}^{C_{21}H_{37}} + 3\,O = \overbrace{COOH\quad \underset{\underset{COOH\;\;CH_2}{\diagup\;\;\diagdown\;\diagup}}{CH\quad CH_2}}^{C_{21}H_{37}}.$$

Diese neue Säure $C_{26}H_{44}O_4$ wurde nun wiederum der Behandlung mit Essigsäureanhydrid und der Destillation im Vakuum unterworfen. Das hierbei entstehende neutrale Destillationsprodukt ist aber kein Keton, sondern nur das Anhydrid der Säure $C_{26}H_{44}O_4$; es besitzt die Formel $C_{26}H_{42}O_3$ und läßt sich leicht in die freie Säure zurückverwandeln Die Säure $C_{26}H_{44}O_4$ verhält sich also wie eine Dicarbonsäure mit der 1,4 oder 1,5 Stellung der Carboxyle. Aus diesen Ergebnissen läßt sich die Gliederzahl von Ring (1) ableiten; denn da die Stellung der Carboxyle in den Säuren $C_{27}H_{46}O_4$ und $C_{26}H_{44}O_4$ nur um eine Einheit verschieden sein kann, muß die Säure $C_{26}H_{44}O_4$ die Carboxyle in 1,5 Stellung, die Säure $C_{27}H_{46}O_4$ sie in 1,6 Stellung enthalten. Der Ring 1 ist also ein Sechsring.

Für die Untersuchung des Ringes 2 schien das Cholesten

$$\overbrace{\begin{matrix} H_2C\diagup & | & | \\ H_2C\diagdown & CH\diagdown & \diagup CH \\ CH_2 & & CH \end{matrix}}^{C_{21}H_{37}}$$

das gegebene Ausgangsmaterial zu sein; eine direkte Oxydation zu der gesuchten Dicarbonsäure gelang indessen nicht; das Cholesten wurde darum über das Nitrocholesten

$$\overbrace{\begin{matrix} H_2C\diagup & | & | \\ H_2C\diagdown & CH\diagdown & \diagup C-NO_2 \\ CH_2 & & CH \end{matrix}}^{C_{21}H_{37}}$$

in das Heterocholestanon

$$\overbrace{\quad\quad\quad\quad\quad\quad}^{C_{21}H_{37}}$$
$$H_2C \quad | \quad\quad |$$
$$H_2C \quad CH \quad CO$$
$$CH_2 \quad CH_2$$

verwandelt und dieses für den Abbau benutzt. Außer dem Heterocholestanon wurde noch auf etwas anderem Wege ein zweites Keton, das die Carbonylgruppe in Ring 2 enthält, dargestellt und zu der gewünschten Dicarbonsäure oxydiert. Wird nämlich Cholesten mit Chromsäureanhydrid oxydiert, entsteht ein ungesättigtes Keton $C_{27}H_{44}O$, das Oxocholesten, das leicht in das entsprechende gesättigte Keton, das Oxocholestan, übergeführt werden kann. Dieses Oxocholestan, das andere Eigenschaften besitzt wie Heterocholestanon, aber wie letzteres bei der Reduktion nach Clemmensen Cholestan liefert, kann sich nur durch die Stellung der Ketogruppe vom Heterocholestanon unterscheiden. Die Stellung der Ketogruppe im Heterocholestanon ist sicher bewiesen; für das Oxocholesten ist es nach mehrfachen Analogien wahrscheinlich, daß die Carbonylgruppe aus einer der doppelten Bindung benachbarten Methylengruppe hervorgegangen ist:

$$\overbrace{\quad\quad\quad\quad\quad\quad}^{C_{20}H_{35}} \quad\quad\quad \overbrace{\quad\quad\quad\quad\quad\quad}^{C_{20}H_{35}}$$
$$H_2C \quad | \quad CH_2 \quad\quad\quad H_2C \quad | \quad CO$$
$$H_2C \quad CH \quad CH \quad \rightarrow \quad H_2C \quad CH \quad CH$$
$$CH_2 \quad CH \quad\quad\quad\quad CH_2 \quad CH$$

Cholesten Oxocholesten

$$\overbrace{\quad\quad\quad\quad\quad\quad}^{C_{20}H_{35}}$$
$$H_2C \quad | \quad CO$$
$$\rightarrow \quad H_2C \quad CH \quad CH_2$$
$$CH \quad CH_2$$

Oxocholestan

Dem Oxocholestan käme dann die obenstehende Formel zu. Daß diese Formulierung wirklich zutrifft, geht aus dem weiteren Abbau des Heterocholestanons und des Oxocholestans hervor. Beide Ketone lassen sich nämlich mit rauchender Salpetersäure ganz leicht aufspalten und geben dabei eine und dieselbe Dicarbon-

säure $C_{27}H_{46}O_4$. Der Vorgang verläuft also bestimmt nach folgendem Schema:

$C_{20}H_{35}$

H_2C CH_2

H_2C CH CO →

CH_2 CH_2

Heterocholestanon

$C_{20}H_{35}$

H_2C $COOH$

H_2C CH $COOH$

CH_2 CH_2

$C_{20}H_{35}$

H_2C CO

← H_2C CH CH_2.

CH_2 CH_2

Oxocholestan

Diese neue durch Aufspaltung des Ringes 2 gewonnene Säure verhält sich nun bei der Destillation ganz anders wie das aus Ring 1 bereitete Isomere; sie liefert nämlich kein Keton, sondern nur das normale Säureanhydrid $C_{27}H_{44}O_3$, das unzersetzt im Vakuum über 270° destillierbar ist und sich leicht in die Ausgangssäure zurückverwandeln läßt. Die Carboxylgruppen dürften hier also in 1,4- oder 1,5 Stellung stehen; der Ring 2 sollte demnach ein Vier- oder Fünfring sein. Ein Vierring läßt sich nun mit den bereits sicher nachgewiesenen Gruppen

$C_{18}H_{35}$

C

H_2C C · · · · CH_2

H_2C CH CH

CHOH CH

nicht mehr konstruieren, es könnte sich also nur um einen *Fünfring* handeln. Es muß indessen hervorgehoben werden, daß diese Folgerungen nur dann zutreffen, wenn bei *komplizierten polycyklischen Gebilden die Leichtigkeit der Ringschließung von ganz ähnlichen Faktoren abhängig ist wie bei den einfacher gebauten Verbindungen*. Die Behauptung, daß das Cholesterin einen hydrierten Indenring enthält, soll daher nur mit einem gewissen Vorbehalt ausgesprochen werden.

Diese Vorsicht ist um so notwendiger, weil bei der Übertragung der Blanc'schen Methode auf Tricarbonsäuren des Cholesterins Ergebnisse erzielt worden sind, die sich mit der Annahme eines hydrierten Indenringes nicht ohne weiteres in Einklang bringen lassen. Auch diese Versuche sollen kurz geschildert werden:

Cholesterin läßt sich über eine Anzahl Zwischenprodukte, von denen nur die Formeln angeführt seien, zu einer Tricarbonsäure $C_{25}H_{42}O_6$ oxydieren:

```
              C18 H35
        ____________________
              C
       H2 C/    \C .... CH2
       H2 C\    /CH\   /CH
            CHOH     CH

       C18 H35                          C18 H35
  ____________________           ____________________
          C                               C
→  H2 C/    \C .... CH2     →     H2 C/    \C ... CH2
 HOOC       /CH\   /CH          HOOC       /C\   /CO
        COOH     CH                    COOH    CH

       C18 H35                          C18 H35
  ____________________           ____________________
          C                               C
→  H2 C/    \C .... CH2     →     H2 C/    \C ...... CH2
 HOOC       /CO    /CO          HOOC      COOH      COOH
        COOH   COOH                  Säure C25 H42 O6 1)
```

Diese Tricarbonsäure $C_{25}H_{42}O_6$ liefert bei der Destillation nach Blanc's Methode eine Ketomonocarbonsäure $C_{24}H_{40}O_3$ [2]) und enthält daher wahrscheinlich zwei Carboxylgruppen in 1,6 oder 1,7-Stellung. Diese Monoketocarbonsäure $C_{24}H_{40}O_3$ wird mittels Chromsäure oder Salpetersäure recht glatt zu einer Tricarbonsäure $C_{24}H_{40}O_6$ aufgespalten:

---

1) In den Originalarbeiten haben alle diese Säuren Formeln mit zwei Atomen Wasserstoff weniger. Ber. D. chem. Ges. **41**, 611 (1908).

2) Windaus, Ber. D. chem. Ges. **45**, 1318 (1912).

$C_{18}H_{35}$

C — C — CH_2 (ring with CH_2, CO), COOH

CH_2

CO

Ketomonocarbonsäure $C_{24}H_{40}O_3$

→

$C_{18}H_{35}$

C — C — CH_2

COOH COOH COOH

Tricarbonsäure $C_{24}H_{40}O_6$

Die Tricarbonsäure $C_{24}H_{40}O_6$ verhält sich nun bei der Destillation nach Blanc gerade so wie die Säure $C_{25}H_{42}O_6$, sie spaltet Kohlendioxyd und Wasser ab und ergibt dabei eine Ketomonocarbonsäure $C_{23}H_{38}O_3$; sie verhält sich also ebenfalls wie eine Säure mit der 1,6- oder 1,7-Stellung zweier Carboxyle.

Dieses letzte Resultat ist unerwartet; wie ein Blick auf die oben stehenden, unter der Annahme eines hydrierten Indenringes abgeleiteten Formeln zeigt, enthält nämlich die Säure $C_{24}H_{40}O_6$ keine Carboxyle in der 1,6- oder 1,7-Stellung. Wie sich dieser Widerspruch auflösen läßt, kann vorläufig aus Mangel an neuem experimentellen Material noch nicht sicher entschieden werden. Wenn die zuletzt erwähnten Experimente allein vorhanden wären, hätte man auf das Vorhandensein eines hydrierten Naphtalinringes im Cholesterin schließen müssen.

Seitenkette im Cholesterinmolekül.

Durch die eben beschriebenen Versuche haben wir über zwei der vier Ringe des Cholesterins Auskuft erhalten; über die beiden anderen hat sich bisher nichts ermitteln lassen; dagegen hat ein energischer oxydativer Abbau noch den Nachweis einer längeren Seitenkette im Molekül gebracht.

Wird die oben erwähnte Säure $C_{25}H_{42}O_6$ mit Chromsäurelösung erhitzt, zerfällt sie in Aceton und eine Tetracarbonsäure[1]) $C_{22}H_{34}O_8$ nach folgender Gleichung:

$$(CH_3)_2CH - CH_2 - C_{18}H_{30}(COOH)_3 + 4O = (CH_3)_2CO + H_2O$$
$$+ HOOC - C_{18}H_{30}(COOH)_3.$$

Die Säure $C_{25}H_{42}O_6$ muß also eine Seitenkette $(CH_3)_2CH - CH_2 -$ enthalten; und dasselbe läßt sich für die Säure $C_{24}H_{40}O_6$ beweisen.

1) Windaus, Berichte D. chem. Ges. **42**, 3772 (1909).

Etwas anders verläuft der Zerfall, wenn man zur Oxydation der Säure $C_{25}H_{42}O_6$ nicht Chromsäure, sondern rauchende Salpetersäure anwendet; hier findet sich unter den Reaktionsprodukten α-Oxy-iso-buttersäure und eine Tetracarbonräure $C_{21}H_{32}O_8$ [1]). Es ist also folgende Reaktion eingetreten:

$$(CH_3)_2\overset{H}{C}-\overset{H_2}{C}-\overset{H_2}{C}-C_{17}H_{28}(COOH)_3 + 6O = (CH_3)_2C(OH)-COOH$$
$$+ H_2O + HOCO-C_{17}H_{28}(COOH)_3$$

und dadurch der Beweis für das Vorhandensein einer fünfgliederigen Seitenkette, des Iso-amylrestes $(CH_3)_2\overset{H}{C}-\overset{H_2}{C}-\overset{H_2}{C}-$ erbracht.

Es ist dann noch ein weiterer Schritt vorwärts geglückt. Bei der Einwirkung sehr verschiedenartiger Oxydationsmittel auf Cholesterin entsteht in kleiner Menge, aber regelmäßig ein wohlriechender Stoff, der, wie eine sorgfältige Prüfung ergeben hat, mit Methyl-iso-hexylketon (Methylheptanon) identisch ist [2]). Leider ist es nicht gelungen das andere Spaltstück, das neben Methylheptanon entstehen muß, zu isolieren und so eine sichere Deutung des Reaktionsverlaufes zu ermöglichen. Am nächsten liegt es gewiß anzunehmen, daß im Cholesterin eine Oktylseitenkette

$$(CH_3)_2\overset{H}{C}-\overset{H_2}{C}-\overset{H_2}{C}-\overset{H_2}{C}-\underset{|}{\overset{H}{C}}-CH_3$$

vorhanden ist, die bei der Oxydation Methyl-iso-hexylketon liefert. Ganz sicher ist diese Annahme zwar nicht, doch läßt sie sich durch mehrere andere Beobachtungen stützen. Wird nämlich Cholesterylchlorid bei gewöhnlichem Druck destilliert, zerfällt es in einen Kohlenwasserstoff mit neunzehn Kohlenstoffatomen und in ein Oktan, dem auch ein Okten beigemischt ist [3]). Es ist wahrscheinlich, daß dieses Okten aus derselben Seitenkette hervorgeht, wie das Methylheptanon. Auch eine Reihe niedrigmolekularer Produkte, die bei energischer Oxydation des Cholesterins mit Salpetersäure entstehen (Dinitro-iso-propan, Bernsteinsäure, Methylbernsteinsäure, α-Methyl-glutarsäure) lassen sich von dieser Oktylseitenkette ableiten.

---

1) Windaus, Berichte D. chem. Ges. **41**, 2558 (1908).

2) Windaus und Resau, Berichte D. chem. Ges. **46**, 1246 (1913).

3) Mauthner und Suida, Monatshefte f. Chemie **17**, 41 (1896).

Konstitution des Cholesterins.

Auf Grund des vorliegenden experimentellen Materials läßt sich also die Formel des Cholesterins wie folgt auflösen:

$$(CH_3)_2\overset{H}{C}-\overset{H_2}{C}-\overset{H_2}{C}-C_{18}H_{24}$$

C
H_2C C CH_2
H_2C CH CH
CHOH CH

Hält man den Nachweis eines Oktylrestes für erbracht, kann die Formel wie folgt geschrieben werden:

$$(CH_3)_2\overset{H}{C}-\overset{H_2}{C}-\overset{H_2}{C}-\overset{H_2}{C}-\overset{H}{C}-CH_3$$

$C_{10}H_{18}$

C
H_2C C CH_2.
H_2C CH CH
CHOH CH

In dem Rest mit 10 Kohlenstoffatomen müssen noch zwei gesättigte Ringe vorhanden sein, die wahrscheinlich einige Kohlenstoffatome mit den Ringen (1) und (2) gemeinsam haben. An sicher ermittelten Einzelheiten fehlt es bisher, doch lassen sich immerhin Strukturformeln ableiten, die ein annähernd richtiges Bild über die Anordnung der Atome in dem großen Cholesterinmolekül geben. Eine der möglichen Konstitutionsformeln ist z. B. die folgende:

CH_2
H_3C-CH CH_2 CH_3
HC $CH-\overset{H_2}{C}-\overset{H}{C}-\overset{H_2}{C}-\overset{H_2}{C}-\overset{H_2}{C}-CH$
H_2C CH CH_3 CH_3
HC CH_2
H_2C C CH_2
H_2C CH CH
CHOH CH

Über die Abhängigkeit der Schwingungszahlen einer Membran von ihrer Begrenzung und über asymptotische Eigenwertverteilung.

Von

R. Courant in Göttingen.

Vorgelegt von C. Runge in der Sitzung vom 28. März 1919.

Die Frage, wie sich die Eigenwerte einer linearen elliptischen partiellen Differentialgleichung für ein Gebiet der unabhängigen Variabeln bei gegebenen Randbedingungen verhalten, wenn das Gebiet sich ändert, ist sowohl vom rein mathematischen, als auch vom physikalischen Standpunkt aus von großem Interesse. Für Differentialgleichungen zweiter Ordnung mit zwei und drei unabhängigen Veränderlichen hat Weyl[1]) die Frage in Angriff genommen und insbesondere das asymptotische Verhalten des $n$ten Eigenwertes mit wachsendem n untersucht. Das Ergebnis dieser Untersuchungen, dessen Kern auf Grund physikalischer Erwägungen vorher als Vermutung von Sommerfeld und von H. A. Lorenz ausgesprochen worden war, lautet etwa bei der Differentialgleichung der schwingenden Membran

(1) $$\Delta u + \lambda u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \lambda u = 0$$

in zwei unabhängigen Variabeln x und y für irgend ein Gebiet G der x-y-Ebene vom Flächeninhalt f unter der Randbedingung $u = 0$

(2) $$\lim_{n=\infty} \frac{\lambda_n}{n} = \frac{4\pi}{f}.$$

1) Math. Annalen, 71, 1912, S. 441—478, Crelles Journal Bd. 141, S. 1—11, sowie ebenda S. 163—181; Crelles Journal Bd. 143, S. 177—202.

Dabei bedeutet λ_n den $n$ten der nach wachsender Größe geordneten Eigenwerte der Differentialgleichung (1) für das Gebiet G und die Randbedingung $u = 0$, physikalisch gesprochen das Quadrat der $n$ten Eigenschwingungszahl einer am Rande eingespannten Membran G. Die asymptotische Eigenwertverteilung ist also von der Gestalt des Gebietes unabhängig und wird lediglich durch den Flächeninhalt bestimmt. Analoge Resultate ergeben sich für andere Randbedingungen, allgemeinere Differentialgleichungen zweiter Ordnung und für drei unabhängige Variabele.

Die Weylschen Untersuchungen beruhen darauf, daß man das Eigenwertproblem der Differentialgleichung in ein Eigenwertproblem für eine Integralgleichung umformt und sodann dieses letztere behandelt.

Man kann nun zu den fraglichen Eigenwertproblemen ohne Vermittelung der Integralgleichungstheorie einen ganz direkten und überraschend einfachen Zugang finden, wenn man als leitenden Gedanken das Prinzip der Variationsrechnung voranstellt, Größen durch ihre Extremaleigenschaften zu charakterisieren. Auf diese Art ergibt sich eine allgemeine Methode, welche sich in einheitlicher Weise auf den ganzen fraglichen Kreis der Eigenwertaufgaben der mathematischen Physik anwenden läßt und deren Ergebnisse die Weylschen in sich schließen. Während ich mir vorbehalte, die vollständige Behandlung der genannten Aufgaben an anderer Stelle zu veröffentlichen, möchte ich hier den Grundgedanken der Methode darlegen, indem ich für das Eigenwertproblem der Differentialgleichung (1) bei der Randbedingung $u = 0$ unter anderen die folgenden Sätze beweise[1]):

Satz 1. *Der $n$te Eigenwert λ_n eines Gebietes G ist kleiner als der $n$te Eigenwert jedes Teilgebietes.*

Satz 2. *Für den $n$ten Eigenwert λ_n des Gebietes G mit dem Flächeninhalt f gilt die asymptotische Beziehung*

$$\lim_{n=\infty} \frac{\lambda_n}{n} = \frac{4\pi}{f}.$$

Satz 1 ist die Verallgemeinerung eines bekannten von H. A. Schwarz herrührenden Satzes über den ersten Eigenwert.

Dabei werde ich die folgende klassische Definition[2]) der Eigen-

---

1) Man vergleiche hierzu besonders die erste Arbeit von Weyl in Crelles Journal, welche dieselben Resultate erzielt, wie die vorliegende Arbeit, und mit welcher die vorliegende Arbeit gewisse Berührungspunkte besitzt.

2) Vergl. etwa H. Weber, Math. Annalen, Band 1, S. 1 ff.

werte λ_n und der zugehörigen Eigenfunktionen u_n zum Ausgangspunkt nehmen. Sind $u_1, u_2, \ldots u_{n-1}$ die $n-1$ ersten Eigenfunktionen der Differentialgleichung (1) für das Gebiet G und die Randbedingungen $u = 0$, λ_n der $n$te Eigenwert, u_n die $n$te Eigenfunktion, ist ferner φ irgend eine der Randbedingung

$$\varphi = 0 \tag{3}$$

genügende, in G sonst stetige und stückweise stetig differenzierbare[1]) Funktion, welche den weiteren Bedingungen

$$\iint\limits_{(G)} \varphi^2 \, dx\, dy = 1, \tag{4}$$

$$\iint\limits_{(G)} \varphi\, u_i \, dx\, dy = 0, \quad (i = 1, \ldots n-1) \tag{5}$$

genügt, so ist

$$D[\varphi] \equiv \iint\limits_{(G)} \left[\left(\frac{\partial \varphi}{\partial x}\right)^2 + \left(\frac{\partial \varphi}{\partial y}\right)^2\right] dx\, dy \geqq \lambda_n, \tag{6}$$

und es gilt

$$D[u_n] = \lambda_n. \tag{7}$$

Da die Funktion $u_n = \varphi$ den Bedingungen (3), (4), (5) genügt, so bedeutet dies: Das Variationsproblem, $D[\varphi]$ zu einem Minimum zu machen, wenn die zur Konkurrenz zuzulassenden Funktionen den obigen Bedingungen genügen, besitzt die $n$te Eigenfunktion als Lösung, und der zugehörige kleinste Wert von $D[\varphi]$ ist gleich dem $n$ten Eigenwert λ_n.

§ 1. Eine Maximum-Minimum-Eigenschaft des $n$ten Eigenwertes.

Unsere Methode beruht auf der Bemerkung, daß man den $n$ten Eigenwert unabhängig von der Kenntnis der zu vorangehenden Eigenwerten gehörigen Eigenfunktionen durch eine Maximum-Minimum-Eigenschaft definieren kann, während die obige Definition den Übelstand aufweist, daß man die Eigenwerte successive gewinnt. Die fragliche Eigenschaft ist die sinngemäße Analogie folgender Eigenschaft quadratischer definiter Formen von

---

1) Stückweise stetig differenzierbar heißt eine stetige Funktion in G, wenn ihre ersten Ableitungen überall in G stetig sind höchstens mit Ausnahme von endlich vielen analytischen Kurvenstücken, längs deren sie endliche Sprünge erleiden dürfen.

m Variabeln oder geometrisch gesprochen, m dimensionaler Ellipsoide:

Ordnet man die Längen der Hauptachsen eines m dimensionalen Ellipsoides in eine fallende Reihe, so ist die $n$te dieser Zahlen gleich dem kleinsten Wert, welchen die größte Hauptachse irgend eines auf dem Ellipsoid gelegenen Schnittellipsoides von $m-n+1$ Dimensionen besitzen kann.

In unserem Falle lautet die fragliche Extremaleigenschaft folgendermaßen:

Seien $v_1, \ldots v_{n-1}$ stetige Funktionen in G, sei $d\{v_1, v_2, \ldots v_{n-1}\}$ die untere Grenze aller Werte, welche das Integral $D[\varphi]$ annehmen kann, wenn φ den Bedingungen (3), (4) und

$$(5^a) \qquad \iint_{(G)} \varphi v_i \, dx\, dy = 0 \qquad (i = 1, \ldots n-1)$$

genügt, im übrigen eine in G stetige und stückweise stetig differenzierbare Funktion ist. Dann ist der $n$te Eigenwert das Maximum aller Werte $d\{v_1, \ldots v_{n-1}\}$, wenn zum Vergleich alle möglichen Funktionensysteme $v_1, \ldots v_{n-1}$ zugelassen werden. Das Maximum wird erreicht für $v_1 = u_1$, $v_2 = u_2$, $\ldots v_{n-1} = u_{n-1}$, $\varphi = u_n$.

Man kann diese Eigenschaft auch folgendermaßen formulieren: Für ein beliebig vorgegebenes System $v_1, v_2, \ldots v_{n-1}$ läßt sich eine den angegebenen Bedingungen genügende Funktion φ finden, sodaß

$$D[\varphi] \leqq \lambda_n$$

wird. Es gibt jedoch ein Funktionensystem, für welches stets

$$D[\varphi] \geqq \lambda_n$$

ist.

Der letzte Teil der Behauptung folgt unmittelbar aus der ursprünglichen Definition der Eigenwerte und Eigenfunktionen, wenn man $v_i = u_i$ setzt. Um den ersten Teil der Behauptung zu erweisen, konstruieren wir zu den vorgegebenen $n-1$ Funktionen $v_1, \ldots v_{n-1}$ eine lineare Kombination

$$\varphi = c_1 u_1 + c_2 u_2 + \ldots + c_n u_n$$

der ersten n Eigenfunktionen, welche den Bedingungen (4), (5ᵃ) genügt. Während die Bedingungen (5ᵃ) $n-1$ lineare homogene Gleichungen für die n Größen c_i ergeben, folgt aus (4)

$$(8) \qquad c_1^2 + c_2^2 + \ldots + c_n^2 = 1,$$

da für zwei verschiedene Eigenfunktionen u_h, u_k die Beziehung

$$\iint_{(G)} u_h u_k \, dx\, dy = 0$$

gilt. Die Bestimmung der c_i ist jedenfalls auf mindestens eine Weise möglich; die so definierte Funktion φ genügt allen oben für φ gestellten Bedingungen, und da für zwei verschiedene Eigenfunktionen u_h, u_k die Beziehung

$$\iint\limits_{(G)} \left[\frac{\partial u_h}{\partial x}\frac{\partial u_k}{\partial x} + \frac{\partial u_h}{\partial y}\frac{\partial u_k}{\partial y}\right] dx\, dy = 0$$

gilt, so folgt

$$D[\varphi] = c_1^2 D[u_1] + c_2^2 D[u_2] + \ldots + c_n^2 D[u_n] = c_1^2\lambda_1 + c_2^2\lambda^2 + \ldots + c_n^2\lambda_n$$

und somit wegen (8) und der Relation $\lambda_i \leqq \lambda_{i+1}$

$$D[\varphi] \leqq \lambda_n$$

wie gezeigt werden sollte. Damit ist die behauptete Extremaleigenschaft des $n$ten Eigenwertes bewiesen.

Es sei hervorgehoben, daß diese Betrachtung auch dann gültig bleibt, wenn das Gebiet G aus einer Anzahl von getrennten Gebieten $G', G'', \ldots$ besteht. Die Eigenwerte des Gebietes sind dann die der Teilgebiete $G', G'', \ldots$ zusammengenommen, die Eigenfunktionen sind die Eigenfunktionen der Einzelgebiete, welche man sich nach außen stets als identisch Null fortgesetzt zu denken hat.

§ 2. Der $n$te Eigenwert als Funktion des Gebietes G.

Aus der eben bewiesenen Tatsache folgt Satz 1 fast unmittelbar. Sei nämlich G' ein Teilgebiet von G, so betrachten wir neben dem eben behandelten Maximum-Minimumproblem, welches den $n$ten Eigenwert λ_n von G charakterisierte, ein zweites Maximum-Minimumproblem, welches aus dem ersten entsteht, indem den zur Konkurrenz zuzulassenden Funktionen φ die weitere Bedingung auferlegt wird, in dem außerhalb G' liegenden Teilgebiet von G identisch zu verschwinden. Bezeichnet $d'\{v_1, v_2, \ldots v_{n-1}\}$ die unter dieser weiteren Voraussetzung sich ergebende untere Grenze der Werte $D[\varphi]$ bei gegebenen $v_1, \ldots v_{n-1}$, so ist offenbar wegen der Verschärfung der Bedingungen für φ

$$d'\{v_1, \ldots v_{n-1}\} \geqq d\{v_1, \ldots v_{n-1}\}.$$

Also ist das Maximum der Werte d' größer oder gleich dem Maximum der Werte d. Nun ist aber das zweite Variationsproblem gerade dasjenige, welches nach § 2 den $n$ten Eigenwert λ'_n von G' charakterisiert. Somit folgt

$$\lambda_n \leqq \lambda'_n, \tag{9}$$

womit der behauptete Satz im wesentlichen bewiesen ist.

Daß in der Relation (9) das Gleichheitszeichen nicht gelten kann, wenn nicht G mit G' zusammenfällt, erkennt man leicht aus der Bemerkung, daß anderenfalls nicht nur G und G', sondern dem eben Bewiesenen zufolge auch alle Teilgebiete von G, welche G' als Teilgebiet enthalten, denselben $n$ten Eigenwert λ_n haben müßten.

Aus der eben bewiesenen Relation (9) ergibt sich — jedenfalls, wenn G von einem Punkte aus konvex ist — fast unmittelbar der Satz:

Der $n$te Eigenwert eines Gebietes ändert sich stetig mit dem Gebiete.

Der $n$te Eigenwert jedes von zwei Gebieten G und G^*, deren Ränder von den Rändern zweier ähnlicher und ähnlich liegender Gebiete B und B^* eingegrenzt werden, liegt nämlich zwischen dem $n$ten Eigenwert des Gebietes B und dem $n$ten Eigenwert des Gebietes B^*, welche ihrerseits bis auf einen Faktor, das Quadrat des Ähnlichkeitsverhältnisses, übereinstimmen. Wenn G und G^* hinreichend nahe beieinander liegen, so können B und B^* so gewählt werden, daß der angegebene Faktor sich von 1 um beliebig wenig unterscheidet, womit die behauptete Stetigkeitseigenschaft erwiesen ist[1]).

Die Betrachtungen, welche zum Beweise von Satz 1 führten, lassen sich leicht so verallgemeinern, daß man zu wesentlich weitergehenden Resultaten gelangt. Man denke sich das Maximum-Minimum-Problem, welches die Eigenwerte des Gebietes G charakterisierte, allgemein dadurch modifiziert, daß man der Funktion φ die weiteren Bedingungen auferlegt, auf vorgegebenen in G liegenden Linien oder Teilgebieten zu verschwinden. Dadurch wird, da der Bereich der zur Konkurrenz zugelassenen Funktionen φ verengert ist, für jedes vorgegebene Funktionensystem $v_1, v_2, \ldots v_{n-1}$ die untere Grenze $d\{v_1, v_2, \ldots v_{n-1}\}$ der Integrale $D[\varphi]$ nach oben verschoben, also auch das Maximum dieser unteren Grenzen. Der Maximum-Minimum-Wert des modifizierten Variationsproblemes ist also nicht kleiner als der des ursprünglichen, d. h. als λ_n. Physikalisch bedeutet dies:

Hält man bei einer rings eingespannten schwingenden Membran G Linienstücke oder Flächenstücke fest, so fällt keine Eigenschwingungszahl, d. h. Grundton und alle Obertöne können Änderungen nur in wachsendem Sinne erfahren.

---

1) Man vergleiche zu dieser Schlußweise die Arbeit des Verfassers in Math. Zeitschrift, Bd. 1, S. 322—323.

Seien etwa $G', G'', G''', \ldots$ eine Anzahl[1]) von Teilgebieten von G, welche keine gemeinsamen inneren Punkte haben und G entweder ganz oder teilweise ausfüllen. Legt man dann den Funktionen φ die Bedingung auf, in G überall auf den Rändern von $G', G'', \ldots$ und in dem außerhalb aller $G^{(i)}$ liegenden Teile von G zu verschwinden, so ergibt sich gemäß § 1 als Lösung des durch die neue Bedingung modifizierten Maximum-Minimum-Problemes der $n$te Eigenwert λ_n^* aus der Reihe der zusammengenommenen Eigenwerte der Gebiete $G', G'', G''', \ldots$, wobei die Eigenwerte dieser Teilgebiete natürlich nach ihrer Vielfachheit zu zählen und die λ_i^* in eine aufsteigende Reihe geordnet zu denken sind. Das Maximum-Minimum wird angenommen für die zu diesem Eigenwert im zugehörigen Gebiet gehörige Eigenfunktion, welche außerhalb als identisch Null fortgesetzt zu denken ist. Physikalisch besagt dies die selbstverständliche Tatsache, daß die Membran nunmehr so schwingt, als ob sie aus den völlig von einander unabhängigen Teilmembranen $G', G'', G''', \ldots$ bestände. Aus der obigen Bemerkung folgt nunmehr unmittelbar

$$\lambda_n \leqq \lambda_n^*, \tag{10}$$

d. h. wir erhalten den folgenden Satz:

Der $n$te Eigenwert eines Gebietes G ist sicher nicht größer als die $n$tgrößte Zahl aus der Gesamtmenge der Eigenwerte einer Anzahl von Teilgebieten von G, welche keine gemeinsamen inneren Punkte haben.

Man kann diese Tatsache auch folgendermaßen ausdrücken:

Die Anzahl der Eigenwerte eines Gebietes G, welche unterhalb einer vorgegebenen Grenze liegen, ist sicher mindestens so groß wie die gesamte Anzahl der unter derselben Grenze liegenden Eigenwerte einer Reihe von Teilgebieten von G, welche keine gemeinsamen inneren Punkte haben.

§ 3. Die asymptotische Eigenwertverteilung[2]).

Aus dem zuletzt bewiesenen Resultat folgt leicht das in Gleichung (2) ausgesprochene asymptotische Eigenwertgesetz. Es bestehe zunächst G aus einer endlichen Anzahl h kongruenter Quadrate von Flächeninhalt $a^2 = \frac{f}{h}$.

---

1) Gemeint ist eine endliche Anzahl, die folgenden Überlegungen gelten aber auch für eine unendliche Anzahl.

2) Man vergleiche zu diesem Paragraphen die parallel laufenden Ausführungen bei Weyl loc. cit.

Nun sind die Eigenfunktionen und Eigenwerte eines solchen Quadrates explizite bekannt und werden durch die Ausdrücke

$$\sin\frac{l\pi x}{a}\sin\frac{k\pi y}{a}, \qquad \frac{\pi^2}{a^2}(l^2+k^2)$$

geliefert. Hieraus ergibt sich in bekannter Weise durch Abzählung der Gitterpunkte für die Anzahl $A'(\lambda)$ der Eigenwerte des Quadrates, welche unterhalb einer Grenze λ liegen,

$$A'(\lambda) \cong \frac{a^2}{4\pi}\lambda, \tag{11}$$

wobei das Zeichen $\cong$ asymptotische Gleichheit, d. h. Gleichheit bis auf einen mit wachsendem λ gegen 1 konvergierenden Faktor bedeutet.

Die Anzahl $A(\lambda)$ der unterhalb einer Schranke λ gelegenen Eigenwerte des Gebietes G ist nun zufolge des letzten Satzes aus § 2 größer oder gleich dem h fachen der Anzahl $A'(\lambda)$, somit

$$A(\lambda) \geqq A'(\lambda)\,.\,h \cong \frac{f}{4\pi}\lambda. \tag{12}$$

Um für diese Anzahl ebenso eine obere Schranke zu finden, ist der nächstliegende Weg, das Gebiet G durch Hinzufügung eines aus $\bar{h}$ Quadraten vom Inhalt a^2 bestehenden Gebietes $\bar{G}$ zu einem großen Quadrat Q vom Inhalt $(f+\bar{f}) = a^2(h+\bar{h})$ zu ergänzen. Für die Anzahlen $A^*(\lambda)$, $\bar{A}(\lambda)$ der Gebiete Q, $\bar{G}$ gelten dann entsprechend zu (11) und (12) die Beziehungen

$$A^*(\lambda) \cong \frac{f+\bar{f}}{4\pi}\lambda, \tag{13}$$

$$\bar{A}(\lambda) \gtrapprox \frac{\bar{f}}{4\pi}\lambda. \tag{14}$$

Addiert man (12) und (14), so folgt

$$A(\lambda)+\bar{A}(\lambda) \gtrapprox \frac{f+\bar{f}}{4\pi}\lambda. \tag{15}$$

Andererseits ist dem letzten Satze von § 2 zufolge

$$A(\lambda)+\bar{A}(\lambda) \leqq A^*(\lambda),$$

also wegen (13)

$$A(\lambda)+\bar{A}(\lambda) \lessapprox \frac{f+\bar{f}}{4\pi}\lambda \tag{16}$$

und diese Beziehungen (15), (16) sind nur dann vereinbar, wenn

gilt:

(17) $$A(\lambda) \cong \frac{f}{4\pi}\lambda,$$

(18) $$\bar{A}(\lambda) \cong \frac{\bar{f}}{4\pi}\lambda.$$

Die Gleichung (17) und die Gleichung (2) drücken aber genau dieselbe Tatsache aus.

Nachdem so (2) für jedes aus kongruenten Quadraten zusammengesetztes Gebiet bewiesen ist, folgt der Beweis für ein beliebiges quadrierbares Gebiet unmittelbar aus der Tatsache, daß man jedes derartige Gebiet mit seinem Flächeninhalt durch solche Quadratgebiete beliebig von innen und außen approximieren kann und daß nach § 2 der $n$te Eigenwert eine stetige, monotone Funktion des Gebietes ist.

§ 4. Vertiefung und Erweiterung der Methode.

Die Überlegung des letzten Paragraphen beruht wesentlich auf der Tatsache, daß man die Eigenwertverteilung der Differentialgleichung für *jedes* Quadrat beherrscht. Die Gewinnung der Gleichung (2) erfordert sowohl die Kenntnis der Eigenwerte des Quadrates Q als auch der Eigenwerte der Quadrate vom Inhalt a^2. Man kann sich jedoch von diesem Übelstande leicht befreien, wenn man den Grundgedanken unserer Methode konsequent verfolgt. Modifiziert man nämlich das Maximum-Minimum-Problem des § 1 jetzt nicht dadurch, daß man den Funktionen φ neue Bedingungen auferlegt, sondern dadurch, daß man die an sie gestellten Bedingungen mildert, so muß umgekehrt das bei der Lösung entstehende Maximum-Minimum nicht größer werden als ursprünglich. Wir nehmen wieder an, daß G aus endlich vielen kongruenten Quadraten besteht, und gestatten der Funktion φ, auf den in G liegenden Seiten dieser Quadrate unstetig zu sein. Dann ergibt sich, wie man unschwer erkennt, als Lösung des Problemes eine Eigenfunktion eines gewissen Teilquadrates, welche aber nunmehr auf den im Inneren von G liegenden Rändern der Randbedingung: Normale Ableitung gleich Null genügt. Da man auch für diese Randbedingung die Eigenwertverteilung der Teilquadrate beherrscht, so ergibt sich leicht die obere Schranke für $A(\lambda)$, ohne daß man auf die Eigenwerte eines umfassenden Quadrates zurückzugreifen braucht.

Durch den hier angegebenen Gedanken, dessen Durchführung

an anderer Stelle erfolgen soll, wird unsere Methode geeignet, in ganz einheitlicher Weise alle hier in Frage kommenden Probleme der Eigenwertverteilung zu lösen. Man erkennt mit ihrer Hilfe leicht, wie wenig die asymptotische Eigenwertverteilung durch die Randbedingungen beeinflußt wird. Was die allgemeineren sich selbst adjungierten Differentialgleichungen anbetrifft, deren Koeffizienten nicht konstant sind, so erhält man die Eigenwertverteilung aus unserer Methode durch die Bemerkung, daß man für hinreichend kleine Quadrate die Eigenwertverteilung bei den verschiedenen in Betracht kommenden Randbedingungen beherrscht, da man hier die Koeffizienten als annähernd konstant ansehen kann.

Zum Schluß sei noch darauf hingewiesen, daß die dargelegte Methode ohne jede Veränderung für eine beliebige Anzahl von unabhängigen Veränderlichen gilt, und daß sie ferner auch für sich selbst adjungierte Differentialgleichungen höherer Ordnung anwendbar bleibt.

Reziprozitätsgesetz und Gauss'sche Summen in quadratischen Zahlkörpern.

Von

E. Hecke in Göttingen.

Vorgelegt in der Sitzung vom 28. März 1919.

In dieser Arbeit werde ich zeigen, wie man jedem algebraischen Zahlkörper Ausdrücke von der Art der Gauß'schen Summen zuordnen kann. Bei ihrer Auswertung ist die wesentliche Schwierigkeit wieder die Bestimmung eines Vorzeichens. Ist diese geleistet, so ergibt sich daraus ein *Beweis des quadratischen Reziprozitätsgesetzes* in dem betr. Zahlkörper. In einem Körper, der nebst sämtlichen konjugierten reell ist, gelingt nun die Ermittelung jenes Vorzeichens auf transzendentem Wege durch Heranziehung der *Thetafunktionen von mehreren Variabeln*, wie es bei den gewöhnlichen Gauß'schen Summen bereits durch Cauchy geschehen ist. Diese Thetafunktionen haben natürlich ebenso spezialisierte Moduln wie die, welche bei der Untersuchung der Dedekind'schen Zetafunktion auftreten.

Auf diese Weise ergibt sich für solche Körper ein Beweis des quadratischen Reziprozitätsgesetzes, der wie alle transzendenten Beweise arithmetischer Sätze gegenüber den mehr algebraischen den Vorzug großer Klarheit und Kürze hat.

Im übrigen wird durch diese Untersuchung auch die *Transformationstheorie der Thetafunktionen* vervollständigt. Das bekannte Problem der Bestimmung einen gewissen dabei auftretenden Konstanten war bisher nur insoweit gelöst, als diese Konstante von H. Weber durch mehrfache Gauß'sche Summen ausgedrückt, und für die Auswertung dieser letzteren ein in jedem einzelnen

Falle erst durchzuführendes rekurrentes Verfahren angegeben wurde[1]).

Ich erwähne noch, daß die unten eingeführten Summen auch bei der *Bestimmung der Klassenzahl* gewisser höherer Körper auftreten, welche den bisherigen Methoden noch nicht zugänglich waren. Hierauf hoffe ich bald zurückzukommen.

In der vorliegenden Arbeit beschränke ich mich auf den Fall des quadratischen Körpers. Die allgemeinen Betrachtungen des § 1 über Gauß'sche Summen sind noch für jeden Körper gültig, dagegen gelten die Sätze von § 2 über Vorzeichenbestimmung und Reziprozitätsgesetz nur für reelle quadratische bezw. allgemeiner für total reelle Körper.

§ 1.

Elementare Eigenschaften der allgemeinen Gauss'schen Summen.

Es sei K ein quadratischer Zahlkörper mit der Grundzahl d, $\sqrt{d}$ sei irgendwie fest gewählt. Ferner bezeichne $S(\omega)$ die Spur der Zahl ω in K, d. h. die Summe der mit ω konjugierten Zahlen. Man sieht sofort, daß

$$S\left(\frac{\omega}{\sqrt{d}}\right)$$

eine ganze rationale Zahl ist, wenn ω ganz ist. Ist ω in gekürzter Form der Idealquotient

$$\omega = \frac{\mathfrak{m}}{\mathfrak{n}}$$

und ist n eine ganze rationale Zahl, welche durch das Ideal $\mathfrak{n}$ teilbar ist, so ist

$$nS\left(\frac{\omega}{\sqrt{d}}\right) \text{ ganz.}$$

Wir wollen $\mathfrak{n}$ als den Nenner von ω bezeichnen.

Bedeutet $\varkappa$ irgend eine ganze Zahl aus K, so ist

$$e^{2\pi i S\left(\frac{\omega\varkappa}{\sqrt{d}}\right)}$$

1) *Krazer*, Lehrbuch der Thetafunktionen pg. 181.

H. Weber, Über mehrfache Gauß'sche Summen. Crelles Journal Bd. 74 (1872) pg. 14.

eine n-te Einheitswurzel, die nur von der Restklasse abhängt, welcher die Zahl $\varkappa$ mod. $\mathfrak{n}$ angehört. Hiernach ist

$$G(\omega) = \sum_{\varrho} e^{2\pi i S\left(\frac{\varrho^2 \omega}{\sqrt{d}}\right)},$$

worin ϱ ein vollständiges Restsystem mod. $\mathfrak{n}$ durchläuft, eine nur von ω abhängige Zahl. Es ist die in Aussicht genommene *Verallgemeinerung der Gauß'schen Summe* auf den Körper K.

Wir verabreden zur Abkürzung: Der Zusatz „ϱ mod. $\mathfrak{n}$" am Zeichen $\sum$ soll bedeuten, daß der Summationsbuchstabe ϱ ein vollständiges Restsystem mod. $\mathfrak{n}$ durchlaufen soll.

Der Zusammenhang mit der Theorie der quadratischen Reste in K wird durch folgenden Satz hergestellt:

Sei der Nenner $\mathfrak{n}$ *von* ω *ein zu 2 primes Ideal und* $\varkappa$ *eine ganze zu* $\mathfrak{n}$ *teilerfremde Zahl. Dann ist*

$$G(\omega\varkappa) = \left(\frac{\varkappa}{\mathfrak{n}}\right) G(\omega). \tag{1}$$

Hierbei bedeutet $\left(\frac{\varkappa}{\mathfrak{n}}\right)$ das quadratische Restsymbol in K.

Der Beweis verläuft im Prinzip wie im rationalen Körper. Sei $\mathfrak{n}$ zunächst ein nicht in 2 aufgehendes Primideal $\mathfrak{p}$. Dann ist

$$G(\omega) = \sum_{\varrho \bmod . \mathfrak{p}} \left(\frac{\varrho}{\mathfrak{p}}\right) e^{2\pi i S\left(\frac{\varrho\omega}{\sqrt{d}}\right)}. \tag{2}$$

Um das einzusehen, ist offenbar nur nötig, zu zeigen, daß

$$\sum_{\varrho \bmod . \mathfrak{p}} e^{2\pi i S\left(\frac{\varrho\omega}{\sqrt{d}}\right)} = 0. \tag{3}$$

Ein vollständiges Restsystem mod. $\mathfrak{p}$ erhält man in der Form

$n_1\varrho_1$, wenn $\mathfrak{p}$ vom ersten Grade;
$n_1\varrho_1 + n_2\varrho_2$, wenn $\mathfrak{p}$ vom zweiten Grade.

Hierbei durchlaufen n_1, n_2 je ein vollständiges Restsystem rationaler Zahlen mod. p, wo p die durch $\mathfrak{p}$ teilbare rationale Primzahl ist. Im ersten Fall ist

$$S\left(\frac{\varrho_1 \omega}{\sqrt{d}}\right) \text{ nicht ganz;}$$

im zweiten Falle

$$S\left(\frac{\varrho_1 \omega}{\sqrt{d}}\right) \text{ und } S\left(\frac{\varrho_2 \omega}{\sqrt{d}}\right) \text{ nicht beide ganz,}$$

weil sonst jede Zahl mit dem Nenner $\mathfrak{p}\sqrt{d}$ eine ganzzahlige Spur hätte, was nicht zutrifft[1]). Summation über n_1, bezw. n_1, n_2 liefert daher in der Tat die Gleichung (3). Daraus folgt sofort (2), wenn

$$\left(\frac{\varkappa}{\mathfrak{p}}\right) = 0 \text{ gesetzt wird, falls } \mathfrak{p} \text{ in } \varkappa \text{ aufgeht,}$$

und damit

$$G(\varkappa\omega) = \left(\frac{\varkappa}{\mathfrak{p}}\right) G(\omega). \tag{4}$$

Sei jetzt $\mathfrak{n}$ eine Potenz eines Primideals $\mathfrak{p}^a (a \geqq 2)$ mit ungerader Norm. Wir führen ein Hilfsideal $\mathfrak{a}$ ein derart, daß

$$\mathfrak{p}\mathfrak{a} = \alpha \text{ eine Zahl,}$$

also

$\omega = \dfrac{\beta}{\alpha^a}$, β eine ganze durch $\mathfrak{a}^a$, aber nicht durch $\mathfrak{p}$ teilbare Zahl.

Man erhält ein vollständiges Restsystem mod. $\mathfrak{p}^a$ in der Form

$$\mu + \varrho\alpha^{a-1}$$

wo μ mod. $\mathfrak{p}^{a-1}$ und ϱ mod. $\mathfrak{p}$ je ein vollständiges Restsystem durchlaufen. Daher ist

$$G\left(\frac{\beta}{\alpha^a}\right) = \sum_{\mu \bmod \mathfrak{p}^{a-1}} \sum_{\varrho \bmod \mathfrak{p}} \exp\left\{2\pi i S\left(\frac{(\mu+\varrho\alpha^{a-1})^2\beta}{\alpha^a\sqrt{d}}\right)\right\}$$

$$= \sum_{\mu \bmod \mathfrak{p}^{a-1}} \exp\left\{2\pi i S\left(\frac{\mu^2\beta}{\alpha^a\sqrt{d}}\right)\right\} \sum_{\varrho \bmod \mathfrak{p}} \exp\left\{2\pi i S\left(\frac{2\mu\varrho\beta}{\alpha^a\sqrt{d}}\right)\right\}.$$

Nach (3) ist die innere Summe 0 oder $N(\mathfrak{p})$ je nach dem Werte von μ. Daraus folgt

$$G\left(\frac{\beta}{\alpha^a}\right) = N(\mathfrak{p})\, G\left(\frac{\beta}{\alpha^{a-2}}\right). \tag{5}$$

Mehrmalige Anwendung dieser Rekursionsformel liefert

---

1) Der Grund für die Richtigkeit dieser Behauptung in beliebigen Zahlkörpern ist natürlich der Dedekind-Landsberg'sche Satz. Vgl. § 1 in meiner Note über die L-Funktionen, diese Nachrichten 1917 pg. 300.

$$G\left(\frac{\beta}{\alpha^a}\right) = N(\mathfrak{p})^{\frac{a}{2}} G(\beta), \text{ wenn } a \text{ gerade.}$$

Also ist diese Summe von β ganz unabhängig. Dagegen wird

$$G\left(\frac{\beta}{\alpha^a}\right) = N(\mathfrak{p})^{\frac{a-1}{2}} G\left(\frac{\beta}{\alpha}\right), \text{ wenn } a \text{ ungerade.}$$

In jedem Falle ist also nach (4)

$$G\left(\frac{\varkappa\beta}{\alpha^a}\right) = \left(\frac{\varkappa}{\mathfrak{p}^a}\right) G\left(\frac{\beta}{\alpha^a}\right), \text{ wenn } \mathfrak{p} \text{ nicht in } \varkappa \text{ aufgeht.} \qquad (6)$$

Endlich sei $\mathfrak{n}$ in seine verschiedenen Primfaktoren zerlegt

$$\mathfrak{n} = \mathfrak{p}_1^{a_1} \mathfrak{p}_2^{a_2} \dots \mathfrak{p}_r^{a_r}.$$

Wir führen Hilfsideale ein, $\mathfrak{a}_1, \mathfrak{a}_2, \dots \mathfrak{a}_r$, die zu $\mathfrak{n}$ prim sind und wofür

$$\mathfrak{p}_1 \mathfrak{a}_1 = \alpha_1, \quad \mathfrak{p}_2 \mathfrak{a}_2 = \alpha_2, \quad \mathfrak{p}_r \mathfrak{a}_r = \alpha_r$$

ganze Zahlen werden. Dann ist

$$\omega = \frac{\beta}{\alpha_1^{a_1} \alpha_2^{a_2} \dots \alpha_r^{a_r}}$$

mit ganzem β, das zu $\mathfrak{n}$ prim ist. Setzen wir dann zur Abkürzung

$$\nu = \alpha_1^{a_1} \dots \alpha_r^{a_r}$$

$$\nu_i = \frac{\nu}{\alpha_i^{a_i}} \qquad i = 1, 2 \dots r,$$

so erhält man offenbar ein vollständiges Restsystem mod. $\mathfrak{n}$ in der Form

$$\varrho = \varrho_1 \nu_1 + \varrho_2 \nu_2 \cdots + \varrho_r \nu_r,$$

wo $\varrho_1, \dots \varrho_r$ unabhängig von einander je ein vollständiges Restsystem mod. $\mathfrak{p}_1^{a_1}$, mod. $\mathfrak{p}_2^{a_2}$ ... durchlaufen. Damit wird

$$e^{2\pi i S\left(\frac{\varrho^2 \beta}{\nu\sqrt{d}}\right)} = e^{2\pi i S\left(\frac{\varrho_1^2 \nu_1^2 \beta}{\nu\sqrt{d}}\right) + \cdots + 2\pi i S\left(\frac{\varrho_r^2 \nu_r^2 \beta}{\nu\sqrt{d}}\right)}$$

$$= e^{2\pi i S\left(\varrho_1^2 \frac{\nu_1 \beta}{\alpha_1^{a_1}\sqrt{d}}\right) + \cdots + 2\pi i S\left(\varrho_r^2 \frac{\nu_r \beta}{\alpha_r^{a_r}\sqrt{d}}\right)}$$

und die ganze Summe G wird ein Produkt von r-Summen ähnlicher Bauart, wo die betr. Nenner aber Primidealpotenzen sind:

$$G\left(\frac{\beta}{\nu}\right) = G\left(\frac{\nu_1\beta}{\alpha_1^{a_1}}\right) \cdot G\left(\frac{\nu_2\beta}{\alpha_2^{a_2}}\right) \cdots G\left(\frac{\nu_r\beta}{\alpha_r^{a_r}}\right) \tag{7}$$

Setzen wir $\varkappa\beta$ an Stelle von β in dieser Gleichung und wenden auf jeden Faktor die Gleichung (6) an, so ergibt sich die Behauptung (1).

Der wichtigste Fall ist der, daß

$$\omega = \frac{1}{\mu\nu},$$

wo μ und ν ganze teilerfremde Zahlen mit ungerader Norm sind. Dann ergibt sich, wenn wir in $G\left(\frac{1}{\mu\nu}\right)$

$$\varrho = \varrho_1\mu + \varrho_2\nu$$

setzen und ϱ_1 ein vollständiges Restsystem mod. ν, ϱ_2 ein solches mod. μ durchlaufen lassen

$$G\left(\frac{1}{\mu\nu}\right) = G\left(\frac{\mu}{\nu}\right) G\left(\frac{\nu}{\mu}\right). \tag{8}$$

Wenden wir unsern Satz auf die beiden Faktoren rechts an, so folgt

$$G\left(\frac{1}{\mu\nu}\right) = \left(\frac{\mu}{\nu}\right)\left(\frac{\nu}{\mu}\right) G\left(\frac{1}{\mu}\right) G\left(\frac{1}{\nu}\right). \tag{9}$$

Die Formeln (8) und (9) sind ersichtlich auch gültig, wenn μ, ν Einheiten sind.

Für den Fall, daß der Nenner von ω zu 2 nicht teilerfremd ist, gilt eine ähnliche Rekursionsformel wie (5).

Sei nämlich

$$2 = \mathfrak{l}^c \mathfrak{l}_1 \qquad (c \geqq 1)$$

wo $\mathfrak{l}$ ein Primideal, das nicht in $\mathfrak{l}_1$ aufgeht. Ferner seien

$$\alpha = \mathfrak{l}\mathfrak{a}, \quad \beta = \mathfrak{b}\mathfrak{a}^a \qquad (a \geqq 1)$$

zwei ganz durch $\mathfrak{a}$ teilbare Zahlen, worin $\mathfrak{a}$, $\mathfrak{b}$ zu 2 prim sind. Der Nenner von

$$\omega = \frac{\beta}{\alpha^a} = \frac{\mathfrak{b}}{\mathfrak{l}^a}$$

ist dann $\mathfrak{l}^a$. Es gilt dann für jede ganze zu 2 prime Zahl $\varkappa$

$$G\left(\frac{\varkappa\beta}{\alpha^a}\right) = N(\mathfrak{l})\, G\left(\frac{\varkappa\beta}{\alpha^{a-2}}\right) \tag{10}$$

solange

$$a-2 \geqq 2c.$$

Zum Beweise bedenken wir wieder, daß

$$\mu + \varrho\alpha^{a-c-1}$$

ein vollständiges Restsystem mod. $\mathfrak{l}^a$ darstellt, wenn

$$\mu \text{ mod. } \mathfrak{l}^{a-c-1}$$
$$\varrho \text{ mod. } \mathfrak{l}^{c+1}$$

je ein vollständiges Restsystem durchlaufen. Dabei ist

$$S\left(\frac{(\mu+\varrho\alpha^{a-c-1})^2}{\alpha^a\sqrt{d}}\varkappa\beta\right) = S\left(\frac{\mu^2\varkappa\beta}{\alpha^a\sqrt{d}}\right) + S\left(\frac{2\mu\varrho\varkappa\beta}{\alpha^{c+1}\sqrt{d}}\right) + S\left(\frac{\varrho^2}{\sqrt{d}}\alpha^{a-2c-2}\varkappa\beta\right).$$

Wenn $a-2 \geqq 2c$ ist das letzte Glied für jedes ϱ ganz. Im mittleren Glied hat das Argument für mindestens ein ϱ den Nenner $\mathfrak{l}$, wenn μ nicht durch $\mathfrak{l}$ teilbar. Ist aber μ durch $\mathfrak{l}$ teilbar, so ist das mittlere S für jedes ϱ ganz. Folglich ist

$$G\left(\frac{\varkappa\beta}{\alpha^a}\right) = \sum_{\substack{\mu \equiv 0\,(\mathfrak{l}) \\ \mu \text{ mod. } \mathfrak{l}^{a-c-1}}} e^{2\pi i S\left(\frac{\mu^2}{\alpha^a\sqrt{d}}\right)\varkappa\beta} N(\mathfrak{l})^{c+1}$$

$$= N(\mathfrak{l})^{c+1} \sum_{\mu \text{ mod. } \mathfrak{l}^{a-c-2}} e^{2\pi i S\left(\frac{\mu^2\varkappa\beta}{\alpha^{a-2}\sqrt{d}}\right)}.$$

Da andrerseits

$$\mu^2 \equiv \mu_1^2\,(\mathfrak{l}^{a-2}),$$

wenn

$$\mu \equiv \mu_1(\mathfrak{l}^{a-2-c}), \quad a-2 \geqq 2c,$$

so ist die obige Summe auch gleich

$$N(\mathfrak{l}) \sum_{\mu \text{ mod. } \mathfrak{l}^{a-2}} e^{2\pi i S\left(\frac{\mu^2\varkappa\beta}{\alpha^{a-2}\sqrt{d}}\right)}$$

da hierin jedes Glied $N(\mathfrak{l})^c$-mal auftritt.

Hieraus folgt die Behauptung (10). Durch mehrmalige Anwendung folgt

$$G\left(\frac{\varkappa\beta}{\alpha^a}\right) = N(\mathfrak{l})^{\frac{a}{2}-c}\, G\left(\frac{\varkappa\beta}{\alpha^{2c}}\right), \text{ wenn } a \text{ gerade;}$$

$$G\left(\frac{\varkappa\beta}{\alpha^a}\right) = N(\mathfrak{l})^{\frac{a-1}{2}-c}\, G\left(\frac{\varkappa\beta}{\alpha^{2c+1}}\right), \text{ wenn } a \text{ ungerade.}$$

Daher

$$G\left(\frac{\varkappa\beta}{\alpha^a}\right) = G\left(\frac{\varkappa_1\beta}{\alpha^a}\right), \quad (a \geqq 2c)$$

wenn

$$\begin{cases} \varkappa \equiv \varkappa_1 \, (\text{mod. } \mathfrak{l}^{2c}) & (a \text{ gerade}) \\ \varkappa \equiv \varkappa_1 \, (\text{mod. } \mathfrak{l}^{2c+1}) & (a \text{ ungerade}). \end{cases} \tag{11}$$

§ 2.

Reziprozität zwischen Gauss'schen Summen.

Der Körper K werde jetzt als reeller quadratischer Körper vorausgesetzt. $\sqrt{d}$ bedeute die positive Wurzel aus der Diskriminante. Um nun eine Beziehung zwischen

$$G\left(\frac{\mu}{\nu}\right) \text{ und } G\left(\frac{\nu}{\mu}\right)$$

zu erhalten, gehen wir von der folgenden Transformationsformel der zugehörigen Thetafunktionen aus: Es werde gesetzt

$$\vartheta(t, t', \varrho; \mathfrak{a}) = \sum_{\mu \equiv 0 \, (\mathfrak{a})} \exp\left\{-\pi\left(t(\mu+\varrho)^2 + t'(\mu'+\varrho')^2\right)\right\}.$$

Dabei bedeuten t, t' Variable mit positiv reellem Teil, $\mathfrak{a}$ ist ein Ideal, ϱ irgend eine Zahl aus K, und die Summation ist über alle Zahlen μ von $\mathfrak{a}$ zu erstrecken. Es gilt nun die Gleichung[1])

$$\vartheta(t, t'; \varrho, \mathfrak{a}) = \frac{1}{\sqrt{d}\,N(\mathfrak{a}^2)\sqrt{t}\sqrt{t'}} \sum_{\nu \equiv 0\left(\frac{1}{\mathfrak{a}\sqrt{d}}\right)} \exp\left\{-\pi\left(\frac{\nu^2}{t} + \frac{\nu'^2}{t'}\right) + 2\pi i S(\varrho\nu)\right\} \tag{12}$$

Auf der rechten Seite bedeuten $\sqrt{t}$, $\sqrt{t'}$ diejenigen in der rechten Halbebene eindeutigen Funktionen, welche für positive t, t' selbst positiv sind.

Lassen wir t, t' gleichzeitig so gegen Null konvergieren, daß die reellen Teile von $\frac{1}{t}$, $\frac{1}{t'}$ ins Unendliche wachsen, so folgt daraus

$$\lim_{t, t' = 0} \sqrt{t}\sqrt{t'}\,\vartheta(t, t'; \varrho, \mathfrak{a}) = \frac{1}{N(\mathfrak{a})\sqrt{d}} \tag{13}$$

also unabhängig von ϱ.

---

1) Vgl. meine Note über die L-Funktionen, diese Nachrichten 1917 pg. 312.

Wir betrachten nun insbesondere die Formel (12) mit den Werten $\varrho = 0$, $\mathfrak{a} = 1$. In diesem Falle ergibt sich

$$\sum_{\mu} \exp\{-\pi(t\mu^2 + t'\mu'^2)\} = \frac{1}{\sqrt{d}\sqrt{t}\sqrt{t'}} \sum_{\mu} \exp\left\{-\frac{\pi}{d}\left(\frac{\mu^2}{t} + \frac{\mu'^2}{t'}\right)\right\}.$$

Die Summation ist beiderseits über alle ganzen Zahlen μ aus K zu erstrecken.

Wir werden nun das Verhalten der beiden Seiten bei Annäherung von t, t' an einen Randpunkt $(i\omega, i\omega')$ untersuchen, und dadurch die gesuchte Beziehung zwischen den Gauß'schen Summen erhalten.

Wir setzen

$$t = w^2 - 2i\frac{\beta}{\alpha\sqrt{d}}, \quad t' = w'^2 + \frac{2i\beta'}{\alpha'\sqrt{d}}$$

w, w' seien dabei positiv, $\frac{\beta}{\alpha}$ habe den Nenner $\mathfrak{a}$. Nachher sollen w, w' gegen Null konvergieren. Es ist dann

$$\vartheta(t, t') = \sum_{\mu} \exp\{-\pi(t\mu^2 + t'\mu'^2)\} = \sum_{\mu} \exp\left\{-\pi(w^2\mu^2 + w'^2\mu'^2) + 2\pi i S\left(\frac{\mu^2\beta}{\alpha\sqrt{d}}\right)\right\}.$$

Hierin ersetzen wir μ durch $\lambda + \varrho$, wo ϱ ein vollständiges Restsystem mod. $\mathfrak{a}$, λ alle Zahlen aus $\mathfrak{a}$ durchläuft; damit wird

$$\vartheta(t, t') = \sum_{\varrho \bmod. \mathfrak{a}} \exp\left\{+2\pi i S\left(\frac{\varrho^2\beta}{\alpha\sqrt{d}}\right)\right\} \sum_{\lambda \equiv 0(\mathfrak{a})} \exp\{-\pi(w^2(\lambda+\varrho)^2 + w'^2(\lambda'+\varrho')^2)\}.$$

Die einzelne innere Summe über λ hat nach Multiplikation mit ww' nach (13) einen von ϱ unabhängigen Grenzwert, wenn w, w' gegen Null konvergiert, und so ergibt sich

$$\lim_{w, w' = 0} ww'\vartheta\left(w^2 - \frac{2i\beta}{\alpha\sqrt{d}},\ w'^2 + \frac{2i\beta'}{\alpha'\sqrt{d}}\right) = \frac{1}{\sqrt{d}\,N(\mathfrak{a})} G\left(\frac{\beta}{\alpha}\right).$$

Analog behandeln wir die rechte Seite. Sei $\mathfrak{b}_1$ der Nenner von $\frac{\alpha}{4\beta}$. Es ist

$$\frac{1}{dt} = \frac{1}{d\left(w^2 - \frac{2i\beta}{\alpha\sqrt{d}}\right)} = \frac{i\alpha}{2\beta\sqrt{d}} + \frac{w^2}{\frac{4\beta^2}{\alpha^2} + \frac{2i\beta\sqrt{d}}{\alpha} w^2}$$

$$\frac{1}{dt'} = \qquad - \frac{i\alpha'}{2\beta'\sqrt{d}} + \frac{w'^2}{\frac{4\beta'^2}{\alpha'^2} - \frac{2i\beta'\sqrt{d}}{\alpha'} w'^2}.$$

Daher

$$\lim_{w,\,w'=0} \sqrt{\frac{w^2 w'^2}{\left(\frac{4\beta^2}{\alpha^2}+\frac{2i\beta\sqrt{d}}{\alpha}w^2\right)\left(\frac{4\beta'^2}{\alpha'^2}-\frac{2i\beta'\sqrt{d}}{\alpha'}w'^2\right)}}\,\vartheta\left(\frac{1}{dt},\frac{1}{dt'}\right)=\frac{1}{\sqrt{d}\,N(\mathfrak{b})}\,G\Big(\frac{-}{4}\cdots$$

d. h.

$$\lim w\,w'\,\vartheta\left(\frac{1}{dt},\frac{1}{dt'}\right)=4\left|\frac{\beta\beta'}{\alpha\alpha'}\right|\frac{1}{\sqrt{d}\,N(\mathfrak{b})}\,G\left(\frac{-\alpha}{4\beta}\right).$$

Multiplizieren wir daher die Transformationsgleichung (12) mit $w\,w'$ und lassen w, w' gegen Null konvergieren, so ergibt sich die gewünschte Relation

$$\frac{G\left(\frac{\beta}{\alpha}\right)}{N(\mathfrak{a})}=4\left|\frac{\beta\beta'}{\alpha\alpha'}\right|\cdot\frac{1}{\sqrt{d}\sqrt{\frac{-2i\beta}{\alpha\sqrt{d}}}\sqrt{\frac{+2i\beta'}{\alpha'\sqrt{d}}}}\,\frac{G\left(\frac{-\alpha}{4\beta}\right)}{N(\mathfrak{b})}.$$

Hierin sind die beiden Wurzeln so zu wählen, daß sie als Grenzwerte von Zahlen mit positiv reellem Teil erscheinen, d. h. ihre Amplitude liegt zwischen $-\frac{\pi}{4}$ und $+\frac{\pi}{4}$, also

$$\sqrt{\frac{-2i\beta}{\alpha\sqrt{d}}}=e^{-\frac{\pi i}{4}\operatorname{sgn}\beta\alpha}\left|\sqrt{\frac{2\beta}{\alpha\sqrt{d}}}\right|$$

$$\sqrt{\frac{2i\beta'}{\alpha'\sqrt{d}}}=e^{\frac{\pi i}{4}\operatorname{sgn}\beta'\alpha'}\left|\sqrt{\frac{2\beta'}{\alpha'\sqrt{d}}}\right|.$$

Also lautet die Gleichung

$$G\left(\frac{\beta}{\alpha}\right)=e^{\frac{\pi i}{4}(\operatorname{sgn}\beta\alpha-\operatorname{sgn}\beta'\alpha')}\,2\left|\sqrt{\frac{\beta\beta'}{\alpha\alpha'}}\right|\frac{N(\mathfrak{a})}{N(\mathfrak{b}_1)}\,G\left(-\frac{\alpha}{4\beta}\right). \qquad (14)$$

§ 3.

Das quadratische Reziprozitätsgesetz.

Nehmen wir in der letzten Formel α und β beide zu 2 prim an. Ist dann $\mathfrak{b}$ der Nenner von $\frac{\alpha}{\beta}$, so ist

$$\mathfrak{b}_1 = 4\mathfrak{b}$$

$$\frac{N(\mathfrak{a})}{N(\mathfrak{b})} = \left|\frac{N(\alpha)}{N(\beta)}\right|$$

und es kommt

$$\frac{G\left(\frac{\beta}{\alpha}\right)}{\sqrt{N(\mathfrak{a})}} = \frac{e^{\frac{\pi i}{4}(\operatorname{sgn}\beta\alpha - \operatorname{sgn}\beta'\alpha')}}{8} \frac{G\left(\frac{-\alpha}{4\beta}\right)}{\sqrt{N(\mathfrak{b})}}.$$

Da β und 4 teilerfremd, so ist nach § 1

$$G\left(\frac{-\alpha}{4\beta}\right) = G\left(\frac{-\alpha 4}{\beta}\right)G\left(\frac{-\alpha\beta}{4}\right) = G\left(\frac{-\alpha}{\beta}\right)G\left(\frac{-\alpha\beta}{4}\right)$$

also

$$\frac{G\left(\frac{\beta}{\alpha}\right)}{\sqrt{N(\mathfrak{a})}} = e^{\frac{\pi i}{4}(\operatorname{sgn}\beta\alpha - \operatorname{sgn}\beta'\alpha')} \frac{G\left(\frac{-\alpha\beta}{4}\right)}{8} \frac{G\left(\frac{-\alpha}{\beta}\right)}{\sqrt{N(\mathfrak{b})}}$$

insbesondere für $\beta = 1$

$$\frac{G\left(\frac{1}{\alpha}\right)}{\left|\sqrt{N(\alpha)}\right|} = e^{\frac{\pi i}{4}(\operatorname{sgn}\alpha - \operatorname{sgn}\alpha')} \frac{G\left(\frac{-\alpha}{4}\right)}{8}. \tag{15}$$

Nehmen wir auch noch $\alpha = \pm 1$, so folgt

$$G\left(\frac{\mp 1}{4}\right) = 8,$$

also allgemein für jede Quadratzahl mit ungerader Norm

$$G\left(\frac{\pm \xi^2}{4}\right) = 8. \tag{16}$$

Weiter folgt, daß für jedes zu 2 prime α

$$G\left(\frac{-\alpha}{4}\right) \neq 0.$$

Wäre nämlich $G\left(\frac{-\alpha}{4}\right) = 0$, so müßte nach (15) auch $G\left(\frac{1}{\alpha}\right) = 0$ sein. Man bestimme alsdann eine zu α prime Zahl β, sodaß

$$\alpha\beta \equiv 1 \pmod{4}.$$

Dann ist nach (15), (16)

$$G\left(\frac{1}{\alpha\beta}\right) \neq 0,$$

aber nach (9)

$$G\left(\frac{1}{\alpha\beta}\right) = \left(\frac{\beta}{\alpha}\right)\left(\frac{\alpha}{\beta}\right) G\left(\frac{1}{\alpha}\right) G\left(\frac{1}{\beta}\right)$$

folglich wäre auch $G\left(\frac{1}{\alpha}\right) \neq 0$, entgegen der Annahme.

Das Reziprozitätsgesetz für zwei teilerfremde Zahlen α, β mit ungerader Norm ist nun eine einfache Folge von (14). Wendet man nämlich diese Gleichung auf

$$G\left(\frac{1}{\alpha}\right),\quad G\left(\frac{1}{\beta}\right),\quad G\left(\frac{1}{\alpha\beta}\right)$$

an und berücksichtigt (9) in § 1, so folgt

$$e^{\frac{\pi i}{4}(\operatorname{sgn}\alpha\beta - \operatorname{sgn}\alpha'\beta')} \frac{G\left(\frac{-\alpha\beta}{4}\right)}{8} = \left(\frac{\beta}{\alpha}\right)\left(\frac{\alpha}{\beta}\right) e^{\frac{\pi i}{4}(\operatorname{sgn}\alpha + \operatorname{sgn}\beta - \operatorname{sgn}\alpha' - \operatorname{sgn}\beta')} \frac{G\left(\frac{-\alpha}{4}\right) G\left(\frac{-\beta}{4}\right)}{8^2}$$

Ist jetzt eine der beiden Zahlen α, β *primär*, d. h. quadratischer Rest mod. 4, so ist nach (16) die eine der beiden G rechts gleich 8, während die andere mit dem G links übereinstimmt und, da sie $\neq 0$ ist, sich heraushebt. Somit wird

$$\left(\frac{\beta}{\alpha}\right)\left(\frac{\alpha}{\beta}\right) = e^{\frac{\pi i}{4}(\operatorname{sgn}\alpha\beta - \operatorname{sgn}\alpha - \operatorname{sgn}\beta - \operatorname{sgn}\alpha'\beta' + \operatorname{sgn}\alpha' + \operatorname{sgn}\beta')}$$

$$= e^{\frac{\pi i}{4}((\operatorname{sgn}\alpha - 1)(\operatorname{sgn}\beta - 1) - (\operatorname{sgn}\alpha' - 1)(\operatorname{sgn}\beta' - 1))}.$$

Da jede der Klammern im Exponenten durch 2 teilbar ist, können wir diese Gleichung auch schreiben

$$\left(\frac{\beta}{\alpha}\right)\left(\frac{\alpha}{\beta}\right) = (-1)^{\frac{\operatorname{sgn}\alpha - 1}{2}\frac{\operatorname{sgn}\beta - 1}{2} + \frac{\operatorname{sgn}\alpha' - 1}{2}\frac{\operatorname{sgn}\beta' - 1}{2}} \tag{17}$$

und das ist das *Reziprozitätsgesetz* für zwei teilerfremde Zahlen ungerader Norm, von denen mindestens eine primär ist.

Hierin ist auch der eine *Ergänzungssatz* enthalten. Es bedeute $\mathfrak{a}$ ein solches zu 2 primes Ideal, zu dem es ein Ideal $\mathfrak{c}$ gibt, sodaß

$$\alpha = \mathfrak{a}\mathfrak{c}^2$$

eine primäre Zahl ist. Bedeutet alsdann β irgend eine total positive Zahl, die ein Idealquadrat oder eine Einheit ist, so ist

$$\left(\frac{\alpha}{\beta}\right) = 1$$

und daher wegen

$$\left(\frac{\beta}{\alpha}\right) = \left(\frac{\beta}{\mathfrak{a}}\right)$$

folgt aus (17)

$$\left(\frac{\beta}{\mathfrak{a}}\right) = +1$$

gültig für jedes total positive Idealquadrat β.

Die Umkehrung dieses ersten Ergänzungssatzes erfordert noch einige algebraische Betrachtungen.

Der *zweite Ergänzungssatz* folgt aus den Formeln am Schluß von § 1 und der Gleichung (14). Es bedeute λ eine ganze Zahl, in der nur die Primfaktoren von 2 aufgehen:

$$\lambda = \mathfrak{l}_1^{a_1}\mathfrak{l}_2^{a_2}.$$

Dann ist für jede zu 2 prime Zahl α

$$G\left(\frac{\lambda}{\alpha}\right) = \left(\frac{\lambda}{\alpha}\right) G\left(\frac{1}{\alpha}\right).$$

Auf jede der beiden Summen G werde die Reziprozitätsformel (14) angewandt:

$$\begin{aligned} \frac{G\left(\frac{\lambda}{\alpha}\right)}{|\sqrt{N(\alpha)}|} &= \frac{e^{\frac{\pi i}{4}(\operatorname{sgn}\lambda\alpha - \operatorname{sgn}\lambda'\alpha')}}{8} \frac{G\left(\frac{-\alpha}{4\lambda}\right)}{|\sqrt{N(\lambda)}|} \\ \frac{G\left(\frac{1}{\alpha}\right)}{|\sqrt{N(\alpha)}|} &= e^{\frac{\pi i}{4}(\operatorname{sgn}\alpha - \operatorname{sgn}\alpha')} \frac{G\left(\frac{-\alpha}{4}\right)}{8}. \end{aligned} \tag{18}$$

Mit Rücksicht auf (16) folgt aus den letzten 3 Gleichungen

$$\left(\frac{\lambda}{\alpha}\right) = e^{+\frac{\pi i}{4}(\operatorname{sgn}\lambda\alpha - \operatorname{sgn}\lambda'\alpha' - \operatorname{sgn}\alpha + \operatorname{sgn}\alpha')} \frac{G\left(-\frac{\alpha}{4\lambda}\right)}{8|\sqrt{N(\lambda)}|}, \tag{19}$$

wenn α eine primäre Zahl ist.

Aus (18) folgt weiter, wenn man $\alpha = \xi^2$ wählt,

$$\frac{G\left(\frac{-\xi^2}{4\lambda}\right)}{8|\sqrt{N(\lambda)}|} = e^{\frac{\pi i}{4}(\operatorname{sgn}\lambda' - \operatorname{sgn}\lambda)}$$

Andrerseits hängt die Gauß'sche Summe in (19) nach (11) nur davon ab, welcher Restklasse α mod. $4\mathfrak{l}_1$, mod. $4\mathfrak{l}_2$, bezw. mod. $4\mathfrak{l}_1\mathfrak{l}_2$, mod. 4 angehört, je nachdem, ob die Exponenten a_1, a_2 in λ gerade oder ungerade sind.

Somit folgt aus (11), (19) das zweite Ergänzungsgesetz

$$\left(\frac{\lambda}{\alpha}\right) = (-1)^{\frac{\operatorname{sgn}\alpha-1}{2}\,\frac{\operatorname{sgn}\lambda-1}{2}+\frac{\operatorname{sgn}\alpha'-1}{2}\,\frac{\operatorname{sgn}\lambda'-1}{2}}$$

wenn

$$\alpha \equiv \xi^2 \begin{cases} \text{mod. } 4 & , a_1, a_2 \text{ gerade} \\ \text{mod. } 4\mathfrak{l}_1 & , a_1 \text{ ungerade, } a_2 \text{ gerade} \\ \text{mod. } 4\mathfrak{l}_2 & , a_1 \text{ gerade, } a_2 \text{ ungerade} \\ \text{mod. } 4\mathfrak{l}_1\mathfrak{l}_2, & a_1, a_2 \text{ ungerade.} \end{cases}$$

Zum Schluß bemerke ich noch, daß jede unserer Summen G eine einfache oder doppelte Gauß'sche Summe im Sinne von Weber ist. Sein Verfahren zur Bestimmung ihres Wertes liefert nun in Verbindung mit unserer Gleichung (1) eine Beziehung zwischen quadratischen Restsymbolen in K und im rationalen Grundkörper. Hieraus fließt ein elementarer, aber etwas umständlicher neuer Beweis des Reziprozitätsgesetzes in K, der von der Theorie der Gauß'schen Summen unabhängig ist. Der Gedanke, welcher ihm zu Grunde liegt, findet sich bereits bei Dirichlet[1]), der ähnlich das Reziprozitätsgesetz in $K(\sqrt{-1})$ aus dem in $K(1)$ folgert[2]).

Göttingen, 21. März 1919.

---

1) *Dirichlet*, Recherches sur les formes quadratiques à coefficients et à indéterminées complexes. Crelles Journal Bd. 24 = Werke Bd. I. pg. 535. (§ 6).

2) Der hiermit angedeutete Beweis ist inzwischen — unabhängig von mir — ausgeführt worden in einer eben erscheinenden Arbeit von *Welmin*, Crelles Journal Bd. 149 (1919) pg. 147. (Zusatz bei der Korrektur).

Zur arithmetischen Theorie der algebraischen Grössen.

Von

Alexander Ostrowski in Göttingen.

Vorgelegt von F. Klein in der Sitzung vom 20. Juni 1919.

In der Kroneckerschen Festschrift „Grundzüge einer arithmetischen Theorie der algebraischen Größen“ [1]) wird bekanntlich die Idealtheorie nicht nur für Zahlkörper, sondern allgemein für Körper entwickelt, in denen ganzzahlige algebraische Funktionen einer oder mehrerer Variabeln vorkommen [2]).

Ist $R = [x_1, \ldots, x_n]$ der Körper aller rationalzahliger Funktionen von $x_1, \ldots, x_n$, so gelten für die ganzen Größen dieses Körpers — für ganze ganzzahlige Polynome in $x_1, \ldots, x_n$ — dieselben Zerlegungsgesetze, wie für ganze rationale Zahlen. Es sei K ein in bezug auf $[x_1 \ldots, x_n]$ algebraischer Körper vom Grade m. Nennt man jede algebraische Funktion der x_i ganz, wenn sie Wurzel einer algebraischen Gleichung ist, deren höchster Koeffizient 1 ist und die übrigen Koeffizienten ganze Größen von $[x_1, \ldots, x_n]$

1) Journal f. d. r. u. a. Math., Bd. 92 Ss. 1—122.

2) Von dieser Theorie hat Kronecker wiederholt Gebrauch gemacht in seinen Untersuchungen über die Theorie der elliptischen Funktionen. Vgl. z. B. die Bemerkung in der Festschrift, S. 122, sowie die Abhandlungen in den Sitzungsberichten der Berliner Akademie, insbesondere die Abhandlung: Zur Theorie der elliptischen Funktionen, Sitzungsberichte vom 29. Juli 1886. Als eine Variable des Grundkörpers wird z. B. oft $\varkappa + \frac{1}{\varkappa}$ benutzt, wo $\varkappa$ der Modul ist. — Mit diesen festen Variablen des Körpers sind nicht zu verwechseln die Unbestimmten, die zur Bildung der Fundamentalform und zur Darstellung der Primteiler des Körpers (bei Weber und König zur Darstellung der Funktionale des Körpers) benutzt werden und in beliebiger Anzahl adjungiert werden können. Natürlich kann es trotzdem oft bequem sein, auch gewisse dieser Unbestimmten vorübergehend als Variabeln des Körpers aufzufassen.

sind, so gelten für die Zerlegung der ganzen Größen von K nicht mehr dieselben einfachen Gesetze, wie in $[x_1 \ldots, x_n]$ selbst. Es kann aber, ganz analog wie in der Idealtheorie der Zahlkörper, durch Einführung von idealen Primfaktoren erreicht werden, daß die Arithmetik des Körpers K vollständig durchsichtig wird [1]). Von der so entstehenden Idealtheorie ist die Idealtheorie der Zahlkörper nur ein sehr spezieller Fall. Während aber für die Idealtheorie der Zahlkörper fast alle wichtigen Fragen allgemeinen Charakters bereits vollständig aufgeklärt sind, hat im allgemeinen Falle, daß R Variabeln enthält, einer der wesentlichsten Punkte bisher noch nicht erledigt werden können.

Die Kummersche Idealtheorie der Kreisteilungskörper benutzt sehr wesentlich die Tatsache, daß alle ganzen Zahlen eines solchen Körpers sich als ganze ganzzahlige Funktionen einer bestimmten ganzen Zahl des Körpers darstellen lassen. Mit Hilfe dieser Tatsache gestaltete sich sowohl die Begründung der Theorie als auch die Beweisführung der meisten Sätze besonders einfach. Bei den allgemeinen endlichen Zahlkörpern K gilt aber diese Tatsache im allgemeinen nicht mehr. Ist $\varDelta$ die Diskriminante von K, θ eine beliebige ganze Größe aus K, $d(\theta) = f^2 \varDelta$ ihre Diskriminante, d. h. die Diskriminante des Systems $(1, \theta, \theta^2, \ldots, \theta^{m-1})$, so läßt sich jede ganze Größe von K in der Form darstellen:

$$\frac{a_0 + a_0 \theta + \cdots + a_{m-1} \theta^{m-1}}{f},$$

wo $a_0, a_1, \ldots, a_{m-1}$ ganze Größen von R sind. Damit aber alle ganzen Größen von K sich als ganzzahlige Polynome in θ darstellen lassen, ist notwendig und hinreichend, daß die Diskriminante $d(\theta)$ von θ gleich der Diskriminante $\varDelta$ von K ist. Und dieselbe Bedingung gilt, wenn R Variabeln $x_1, \ldots, x_n$ enthält. Sie ist aber im allgemeinen nicht befriedigt. Denn es seien

$$\mu_1, \mu_2, \ldots, \mu_r$$

Elemente der Fundamentalbasis für die ganzon Größen von K, sodaß jede ganze Größe von K sich in der Form

$$(1) \qquad \mathrm{M} = v_1 \mu_1 + v_2 \mu_2 + \cdots + v_r \mu_r$$

darstellen läßt, wenn man für die v_i geeignete ganze Größen von

1) Eine ausführliche und im allgemeinen korrekte Darstellung dieser Theorie gibt J. König in seinem Buche: Einleitung in die allgemeine Theorie der algebraischen Größen. Leipzig, Teubner 1903.

R wählt[1]). Der Ausdruck (1) genügt für unbestimmte v einer Gleichung m-ten Grades

$$(2) \qquad \Phi(\xi) = 0,$$

deren höchster Koeffizient 1 ist und die übrigen ganzzahlige Polynome in den v_i und x_i sind. Aus dieser Fundamentalgleichung, der die Fundamentalform M genügt, läßt sich für jede primitive ganze Größe von K die irreduzible Gleichung m-ten Grades herleiten, der sie in R genügt, indem man für die v_i die entsprechenden ganzen Größen von R einsetzt. Ebenso erhält man aus der Fundamentaldiskriminante $\varDelta(v_i)$ der Fundamentalgleichung (2) die Diskriminante jeder ganzen Größe $\bar{v}_1\mu_1 + \cdots + \bar{v}_r\mu_r$ von K, wenn man für die v_i die Größen $\bar{v}_i$ von R einsetzt. Für $\varDelta(v_i)$ gilt nun die Gleichung

$$\varDelta(v_i) = \varDelta\bar{\varDelta}^2(v_i),$$

wo $\bar{\varDelta}(v_i)$ wieder ein ganzzahliges Polynom in den v_i und x_i ist. Soll nun die Diskriminante einer ganzen Größe θ von K gleich $\varDelta$ sein, so befriedigen die entsprechenden Werte der v_i die Diophantische Gleichung

$$\bar{\varDelta}(v_i) = \pm 1,$$

was natürlich in den seltensten Fällen möglich ist.

Dieser Umstand hat Dedekind gezwungen, die Beweise einer Reihe von wichtigen Sätzen seiner Theorie auf Umwegen zu führen.

Hier hat nun Kronecker den kühnen Gedanken gefaßt, anstatt von irgend einer bestimmten primitiven ganzen Größe von K, von der Fundamentalform M bei unbestimmten v_i auszugehen. Jede ganze Größe von K läßt sich durch M in der Form ausdrücken:

$$\frac{a_0 + a_1 \mathrm{M} + \cdots + a_{m-1}\mathrm{M}^{m-1}}{\bar{\varDelta}(v_i)},$$

wo $a_0, a_1, \ldots, a_{m-1}$ ganze ganzzahlige Polynome in den v_i und x_i sind. Ist nun der größte gemeinsame Teiler aller ößen von R, die in $\bar{\varDelta}(v_i)$ als Koeffizienten der Potenzprodukte der v_i auftreten, gleich $\varDelta$, d. h., ist $\bar{\varDelta}(v_i)$ primitiv in bezug auf die v_i —, so läßt sich bei den meisten Betrachtungen ganz ebenso schließen, als wenn M eine ganze Größe des Körpers und $\bar{\varDelta}(v_i)$ gleich 1 wäre. Auf diese Weise

Vgl. Kronecker, Festschrift, Ss. 16—19, König, Ss. 495 ff.

ergeben sich dann die schwierigsten Diskriminantensätze von Dedekind und darüber hinaus die Diskriminantensätze der Hilbertschen Theorie der Relativkörper ohne jede Schwierigkeit[1]).

Die Voraussetzung für diese Kroneckersche Theorie der Fundamentalgleichung ist aber, daß $\bar{\varDelta}(v_i)$ primitiv ist. Man hat für $\bar{\varDelta}(v_i)$ jedenfalls die Darstellung

$$\bar{\varDelta}(v_i) = \bar{\bar{\varDelta}}(v_i)\,d,$$

wo das ganzzahlige Polynom d der x_i (beziehungsweise die ganze Zahl d, falls R keine Variabeln enthält) der größte gemeinsame Teiler aller Koeffizienten der Potenzprodukte der v_i in $\bar{\varDelta}(v)$ und $\bar{\bar{\varDelta}}(v_i)$ primitiv in bezug auf die v_i ist. Und es ist zu beweisen, daß $d = 1$ ist.

Dieser Nachweis ist Kronecker nicht gelungen. Er konnte nur beweisen, daß $\bar{d}$ in jedem Falle nur eine natürliche Zahl und nur durch Primzahlfaktoren von $\varDelta$ aus der Reihe $2, 3, \ldots, m-2$ teilbar sein kann. „Ob in der Wirklichkeit solche überflüssige Zahlenteiler in der Diskriminante der Fundamentalgleichung vorkommen können", sagt Kronecker[2]), „habe ich noch nicht ermitteln können; ich habe mich vergeblich bemüht, ein Beispiel dafür aufzufinden, habe aber ebensowenig vermocht, das Gegenteil zu beweisen". Kronecker beweist nur noch[3]), daß $\bar{d}$ stets gleich 1 ist, falls K ein Galoischer Körper in bezug auf R ist.

Für den einfachsten Fall, daß R keine Variabeln enthält, haben 12 Jahre später fast gleichzeitig Herr Hensel und Herr Mertens den Beweis geführt, daß $d = 1$ ist, und damit erst die Kroneckersche Theorie der Fundamentalgleichung vollständig begründet[4]). Im allgemeinen Falle dagegen, daß R Variabeln ent-

---

1) Vgl. Weber, Lehrbuch der Algebra, Bd. II, Ss. 630—671, Hilbert, Bericht über die Theorie der algebraischen Zahlkörper, Jahresbericht d. D. M.-V., Bd. IV, S. 194—201.

2) Festschrift, S. 113.

3) Festschrift, S. 112.

4) K. Hensel, Journal f. d. r. u. a. Math., Bd. 113, S. 61—83, (1894) (Dat. Oktober 1893, herausgegeben den 20. Januar 1894), F. Mertens, Wiener Ber. CIII, S. 5—40 (1894). (Eingereicht in der Sitzung vom 4. Januar 1894, herausgegeben im Januar-Februar Heft). Während die Überlegungen bei Kronecker und Hensel auf der Existenz der Primidealteiler beruhen und daher die so entwickelte Theorie der Fundamentalgleichung nur für den weiteren Ausbau der Idealtheorie benutzt werden kann, war es Mertens gelungen, auch die Primidealteiler mit Hilfe der Primfaktoren der linken Seite der Fundamentalgleichung zu definieren, und dadurch eine vollständige Begründung der Idealtheorie für Zahlkörper zu geben, in der alle früheren Versuche, diese Begründung auf die Theorie der höheren Kongruenzen zu stützen, ihren Abschluß fanden.

hält, ist diese Frage bis heute offen geblieben[1]). Auch in der neueren Darstellung der Kroneckerschen Theorie im Buche von J. König: „Einleitung in die allgemeine Theorie der algebraischen Größen“ mußte die Frage nach der Existenz von Primzahlen, die in d vorkommen, unerledigt bleiben. König gibt nur eine Methode an, um solche „irreguläre“ Primzahlen, falls sie existieren, in Primidealfaktoren zu zerlegen. Allerdings beruht diese Methode auf einem, wie wir sehen werden, unrichtigen Hilfssatz.

Die von mir durchgeführte Untersuchung, über deren Resultate und Methoden im Folgenden berichtet werden soll, hat zum Ergebnis geführt, daß, sobald der Körper R Variabeln enthält, irreguläre Primzahlen wirklich vorkommen können — auch in Galoischen Körpern —, sodaß die Kroneckersche Theorie in diesem allgemeinsten Falle nicht zu halten[2]), und insbesondere das oben angeführte Kroneckersche Resultat über Galoische Körper falsch ist. In der Tat enthält der von Kronecker flüchtig skizzierte Beweis dieses letzten Resultats ein merkwürdiges Versehen. — Nachdem ich im Folgenden zuerst ein Beispiel ausführlich behandelt habe, gehe ich auf die Verhältnisse, unter denen irreguläre Primzahlen vorkommen können, genauer ein. Da die Übertragung der gewöhnlichen Beweise der Hilbertschen Diskriminantensätze in der Theorie der Relativkörper durch das Auftreten der irregulären Primzahlen unmöglich wird — alle bisher veröffentlichten Beweise beruhen auf der Theorie der Fundamentalgleichung — gebe ich sodann eine neue in allen Fällen gültige Beweismethode an. — Ich beschränke mich zumeist auf die kurze Angabe der Resultate und Skizzierung des Beweisganges. Die ausführlichen Beweise sollen an einem anderen Orte veröffentlicht werden. Nur auf den Beweis des am Schlusse der Abhandlung ausgesprochenen Satzes gehe ich ausführlich ein.

Der einfachste Körper K, in dem eine Primzahl — 2 — irregulär ist, entsteht aus dem Körper $R = [x_1, x_2]$ aller rationaler Funktionen mit rationalen Koeffizienten von x_1 und x_2 durch Adjunktion von $\sqrt{x_1}$ und $\sqrt{x_2}$. Dies ist zugleich ein Galoischer Körper. —

Wir bezeichnen $\sqrt{x_1}$ und $\sqrt{x_2}$ durch ξ_1 bzw. ξ_2 und behaupten, daß das System

---

1) König, S. 532. Vgl. auch Enc. d. sc. math. I **10** S. 334.

2) Daher beruht die von Herrn Hensel in seiner zitierten Abhandlung (S. 76) ausgesprochene Behauptung, der Satz über die Nichtexistenz irregulärer Primzahlen gelte auch im allgemeinen von Kronecker behandelten Falle und ließe sich auch für diesen Fall auf dem von Herrn Hensel eingeschlagenen Wege sehr leicht beweisen, offenbar auf einem Irrtum.

(3) $$1,\ \xi_1,\ \xi_2,\ \xi_1\xi_2$$

eine Fundamentalbasis für die ganzen Größen von K in bezug auf R bilden. Die Diskriminante des Systems (3) ist

$$|1,\ \xi_1,\ \xi_2,\ \xi_1\xi_2|^2 = |1,\ \xi_1|^4\,|1,\ \xi_2|^4 = 2^8 x_1^2 x_2^2.$$

Daher gilt für jede ganze Größe g von K eine Gleichung von der Form

(4) $$2^8 x_1^2 x_2^2 g = a_0(x_1,\ x_2) + a_1(x_1,\ x_2)\,\xi_1 + a_2(x_1,\ x_2)\,\xi_2 + a_3(x_1,\ x_2)\,\xi_1\xi_2,$$

wo a_i ganzzahlige Polynome in $x_1,\ x_2$ sind. Und es ist zu beweisen, daß alle a_i durch $2^8 x_1^2 x_2^2$ teilbar sind. Nun folgt aber aus der obigen Gleichung, wenn wir $x_1 = y_1^2$, $x_2 = y_2^2$ setzen, und durch $\bar{g}$ das aus g entstehende ganzzahlige Polynom in $y_1,\ y_2$ bezeichnen:

$$2^8 y_1^4 y_2^4 \bar{g} = a_0(y_1^2,\ y_2^2) + a_1(y_1^2,\ y_2^2)\,y_1 + a_2(y_1^2,\ y_2^2)\,y_2 + a_3(y_1^2,\ y_2^2)\,y_1 y_2.$$

Hier können keine zwei der vier Bestandteile der rechten Seite auch nur ein Glied gemeinsam haben. Daher ist jeder dieser Bestandteile einzeln durch $2^8 y_1^4 y_2^4$ teilbar. Ist aber etwa $a_1(y_1^2,\ y_2^2)\,y_1$ durch $2^8 y_1^4,\ y_2^4$ teilbar, so folgt hieraus, daß $a_1(x_1,\ x_2)$ durch $2^8 x_1^2 x_2^2$ teilbar ist. Und dasselbe gilt für alle a_i. Daher erhalten wir aus (4) die Gleichung

$$y = \bar{a}_0(x_1,\ x_2) + \bar{a}_1(x_1,\ x_2)\,\xi_1 + \bar{a}_2(x_1,\ x_2)\,\xi_2 + \bar{a}_3(x_1,\ x_2)\,\xi_1\xi_2,$$

in der $\bar{a}_i$ ganzzahlige Polynome in $x_1,\ x_2$ sind. W. z. b. w.

Daher ist die Diskriminante $\varDelta$ des Körpers K gleich $2^8 x_1^2 x_2^2$ und die Fundamentalform M ist $v_1 + v_2\xi_1 + v_3\xi_2 + v_4\xi_1\xi_2$. Um den Zahlenteiler der Fundamentaldiskriminante zu erhalten, bilden wir das Produkt $(\mathrm{M}-\mathrm{M}')(\mathrm{M}-\mathrm{M}'')(\mathrm{M}-\mathrm{M}''')$, dessen Norm die Fundamentaldiskriminante ist. Dieses Produkt ist aber gleich

$$(2v_2\xi_1 + 2v_4\xi_1\xi_2)(2v_3\xi_2 + 2v_4\xi_1\xi_2)(2v_2\xi_1 + 2v_3\xi_2),$$

also jedenfalls durch 2^3 teilbar. Seine Norm ist also jedenfalls durch $2^{3\cdot 4} = 2^{12}$ teilbar. Und da $\varDelta$ nur durch 2^8 teilbar ist, kommt 2 in d wenigstens in der zweiten Potenz vor, ist also irregulär. — Es ist auch nicht schwer zu zeigen, daß die Fundamentaldiskriminante genau durch die zwölfte Potenz von 2 teilbar ist, doch dies ist ohne Belang. — Für die Zerlegung der linken Seite der Fundamentalgleichung mod. 2 erhalten wir aus $\mathrm{M} \equiv \mathrm{M}' \equiv \mathrm{M}'' \equiv \mathrm{M}'''$ (mod. 2)

$$\Phi(z) = (z-\mathrm{M})(z-\mathrm{M}')(z-\mathrm{M}'')(z-\mathrm{M}''')$$
$$\equiv (z^2-\mathrm{M}^2)^2 \equiv (z^2 - v_1^2 - v_2^2 x_1 - v_3^2 x_2 - v_4^2 x_1 x_2) \pmod{2}$$

Wir bemerken noch, daß 2 in K ein Primidealteiler bleibt. Denn sonst müßte die Norm einer ganzen Größe

$$a_0(x_1, x_2) + a_1(x_1, x_2)\,\xi_1 + a_2(x_1, x_2)\,\xi_2 + a_3(x_1, x_2)\,\xi_1\,\xi_2$$

durch 2 teilbar sein können, ohne daß alle Polynome $a_i(x_1, x_2)$ durch 2 teilbar wären. Dies ist aber unmöglich, da diese Norm modulo 2 dem Ausdruck kongruent ist:

$$a_0(x_1^4, x_2^4) + a_1(x_1^4, x_2^4)\,x_1^2 + a_2(x_1^4, x_2^4)\,x_2^2 + a_3(x_1^4, x_2^4)\,x_1^2 x_2^2,$$

dessen vier Bestandteile paarweise keine Glieder gemein haben. Da 2 andererseits ein Diskriminantenteiler ist, so erleidet also in diesem Falle der bekannte Satz[1]), daß jeder Diskriminantenteiler wenigstens einen Primidealteiler mehrfach enthält, eine Ausnahme. Auf solche Ausnahmefälle war bereits Kronecker in seinen Untersuchungen über Multiplikationsgleichungen aufmerksam geworden. An einer anscheinend unbeachtet gebliebenen[2]) Stelle seiner Abhandlung: Zur Theorie der elliptischen Funktionen[3]) bemerkt Kronecker, daß die linke Seite einer Multiplikationsgleichung $T(z) = 0$ nach einem gewissen Primzahlmodul n die folgende Form annimmt, in der θ eine Variable ist:

$$T(z) \equiv z(z^n - \theta) \qquad (\text{mod. } n),$$

die Primzahl n aber im entsprechenden Körper ein Produkt von zwei verschiedenen Primidealteilern μ und μ' wird, obwohl die Diskriminante von $T(z)$ durch n teilbar ist. — Allerdings bedarf die Kroneckersche Behandlung dieses Beispiels insofern einer Ergänzung, als Kronecker an angegebener Stelle nur beweist, daß die Diskriminante der speziellen Gleichung $T(z) = 0$ durch n teilbar ist, nicht aber, daß dies für die Körperdiskriminante, oder, was auf dasselbe hinausläuft, für die Fundamentaldiskriminante der Fall ist. Man kann aber aus der im folgenden entwickelten Theorie leicht folgern, daß die Körperdiskriminante jedesmal durch eine Primzahl p teilbar ist, falls die linke Seite wenigstens einer den Körper definierenden irreduziblen Gleichung modulo p einen Teiler besitzt, der einer rationalen Funktion von z^p kongruent ist. Der einfachste Körper, in dem eine solche Singularität auftritt, entsteht aus $[x]$ durch Adjunktion von $\sqrt{x}$ [4]).

1) Festschrift, S. 62, 117. König, S. 539.

2) In der gesamten einschlägigen Literatur wird der angeführte Satz in der ursprünglichen Fassung der Festschrift zitiert.

3) Berliner Sitzungsberichte vom 29. Mai 1886, S. 767—768.

4) In diesem Körper ist dabei 2 keine irreguläre Primzahl. Die ganze

Um allgemein[1]) untersuchen zu können, wann eine Primzahl p irregulär ist, knüpfen wir an an die Zerlegung der linken Seite der Fundamentalgleichung modulo p in irreduzible Faktoren

$$(5) \qquad \Phi(z) \equiv F_1(z)^{a_1} F_2(z)^{a_2} \dots F_s(z)^{a_s} \qquad (\text{mod. } p),$$

wo $F_i(z)$ vom Grade b_i in bezug auf z ist. Dann entspricht jedem irreduziblen Teiler $F_i(z)$ ein Primidealteiler $\mathfrak{p}_i$ von p so, daß $F_i(\text{M})$ durch $\mathfrak{p}_i$, dagegen durch keinen von $\mathfrak{p}_i$ verschiedenen Primidealteiler von p teilbar ist[2]), und die Zerlegung von p in Primidealteiler lautet

$$p = \mathfrak{p}_1^{e_1} \mathfrak{p}_2^{e_2} \dots \mathfrak{p}_s^{e_s}.$$

Wir bezeichnen allgemein den Grad von $\mathfrak{p}_i$ durch g_i, so daß die Norm $N\mathfrak{p}_i$ von $\mathfrak{p}_i$ gleich p^{g_i} ist. Im allgemeinen ist nun für jedes $\mathfrak{p}_i$ $a_i = e_i$, $b_i = g_i$ und der größte gemeinsame Idealteiler von p und $F_i(\text{M})$ gleich $\mathfrak{p}_i$. Und eine Primzahl ist, wie man leicht einsieht, dann und nur dann irregulär, wenn für wenigstens einen Primidealteiler von ihr dies nicht erfüllt ist.

Wir bemerken nun vor allem, daß jedenfalls stets $a_i b_i = e_i g_i$ ist. Denn aus $\Phi(\text{M}) = 0$ folgt wegen (5), daß der größte gemeinsame Idealteiler von p und $F_i(\text{M})^{a_i}$ gleich $\mathfrak{p}_i^{e_i}$ ist. Und da $F_i(\text{M})^{a_i}$ homogen vom Grade $a_i b_i > 0$ in den Unbestimmten v_i ist, ist der größte gemeinsame Teiler der Koeffizienten aller Potenzprodukte der v_i in der Norm von $F_i(\text{M})^{a_i} + p$ gleich der Norm $N_i = p^{e_i g_i}$ von $\mathfrak{p}_i^{e_i}$. Da aber $\Phi(z)$ die Form hat

Entwicklung in der Festschrift beruht nämlich bei Kronecker auf der Voraussetzung, daß die betreffenden Primzahlen nicht irregulär sind. Erst König glaubte den Satz auch auf irreguläre Primzahlen — falls solche existieren — ausdehnen zu können (König, S. 542). Am zuletzt angegebenen Orte beweist Kronecker, daß eine Primzahl p dann und nur dann Diskriminantenteiler ist, wenn bei der Zerlegung der linken Seite der Fundamentalgleichung modulo p in irreduzible Faktoren entweder ein Faktor mehrfach auftritt, oder wenigstens ein Faktor modulo p einer in z^p rationalen Funktion kongruent ist. Mit Hilfe unserer weiteren Entwicklungen können wir hieraus folgern: Eine Primzahl p ist dann und nur dann Diskriminantenteiler, wenn sie gewisse Primidealfaktoren mehrfach enthält oder wenn wenigstens ein irreduzibler Faktor der linken Seite der Fundamentalgleichung modulo p einer in z^p rationalen Funktion kongruent ist.

1) Die folgenden Resultate gelten zum größten Teil auch dann, wenn der zu Grunde gelegte Körper R nicht nur $[x_1, \dots x_n]$, sondern eine endliche algebraische Erweiterung dieses Körpers ist, z. B. die Gesamtheit aller rationalen Funktionen der x_i mit Koeffizienten aus einem endlichen Zahlkörper.

2) Vgl. etwa König, S. 540—544 Mitte, wo der Satz im wesentlichen richtig bewiesen wird, bis auf leicht zu verbessernde Einzelheiten.

$$g(z)\,[F_i(z)^{a_i}+p]+p\,M(z),$$

ist die Norm von $F_i(\mathrm{M})^{a_i}+p$, als Resultante von $F_i(z)^{a_i}+p$ und $\Phi(z)$, gleich der Resultante von $F_i(z)^{a_i}+p$ und $pM(z)$, also durch $p^{a_i b_i}$ teilbar, da $a_i b_i$ der Grad von $\mathrm{F}_i(z)^{a_i}+p$ ist. Da nun einerseits $\prod\limits_i N_i = N(p) = p^m$ ist, andererseits aber auch $\prod\limits_i p^{a_i b_i} = p^m$ ist, muß $a_i b_i = g_i e_i$ sein.

Ist nun für einen Primidealteiler $\mathfrak{p}_i$ von p $a_i \neq e_i$, daher auch $b_i \neq g_i$, so nennen wir $\mathfrak{p}_i$ irregulär von der ersten Art, ist der größte gemeinsame Teiler von p und $F_i(\mathrm{M})$ durch $\mathfrak{p}_i^2$ teilbar, so soll $\mathfrak{p}_i$ irregulär von der zweiten Art heißen. Beide Arten der Irregularität können auch zugleich bei einem und demselben Primidealteiler vorkommen.

Wir gehen zuerst auf die Irregularität erster Art ein. (Hierher gehört das im Obigen ausführlich behandelte Beispiel). Offenbar kann man b_i auch definieren, als den Grad der modulo p irreduziblen Gleichung niedrigsten Grades, der die Fundamentalform M modulo $\mathfrak{p}_i$ genügt. Und es handelt sich jetzt hauptsächlich darum, eine begriffliche Bedeutung für g_i zu finden. Dies gelingt auf die folgende Weise:

Läßt sich eine ganze oder gebrochene Größe von K so in der Form $\frac{M}{N}$ darstellen, daß M und N ganz sind, N zu $\mathfrak{p}_i$ teilerfremd ist, so nennen wir sie ganz in bezug auf $\mathfrak{p}_i$. Ist dabei M durch $\mathfrak{p}_i$ teilbar, so nennen wir sie durch $\mathfrak{p}_i$ teilbar. Betrachten wir nun zwei Größen von K, die modulo $\mathfrak{p}_i$ kongruent sind, deren Differenz also durch $\mathfrak{p}_i$ teilbar ist, als einander gleich, so bilden die in bezug auf $\mathfrak{p}_i$ ganzen Größen von K einen gewissen Kongruenzkörper K. K ist eine algebraische Erweiterung des Kongruenzkörpers P, der aus R entsteht, wenn wir, ganz analog wie für K, die Größen von R, deren Differenz durch p teilbar ist, als gleich ansehen[1]). Der Grad des Körpers K in bezug auf P (d. h. die Maximalzahl von Größen von K, die in bezug auf P linear unabhängig sind) ist nun gleich

1) Diese bekannte und naheliegende Verallgemeinerung des Teilbarkeits- und Kongruenzbegriffs hat den Vorzug, daß sie auch für gebrochene Größen gilt. Sie wird unentbehrlich, wenn in R auch Variabeln vorkommen. Die von König eingeführte Begriffsbildung des Äquivalenzmoduls läuft im wesentlichen auf eine Benutzung eines solchen Kongruenzbegriffes hinaus. Im Folgenden sind stets diese Begriffsbildungen zugrunde gelegt, sofern sich nicht aus dem Zusammenhang das Gegenteil ergibt.

g_i, wenn p^{g_i} die Norm von $\mathfrak{p}_i$ ist. Diese fundamentale Tatsache läßt sich ganz analog beweisen, wie der entsprechende Satz in der Theorie der Zahlkörper[1]). Man kann nun zeigen, daß der Grad von $F_i(z)$ dem größten Grade gleich ist, den eine Größe von K in bezug auf P haben kann. Gibt es daher in K primitive Größen in bezug auf P, so ist $b_i = g_i$. In derartigen Körpern aber, in denen die Vielfachen einer natürlichen Primzahl p gleich 0 angenommen werden, brauchen jedoch nach den Untersuchungen von Steinitz[2]) endliche algebraische Erweiterungen keine primitiven Größen zu besitzen. Hat aber K keine primitive Größe in bezug auf P, so ist $b_i < g_i$, $a_i > e_i$ und $\mathfrak{p}_i$ irregulär von der ersten Art. Damit ist also die Untersuchung der Irregularität erster Art auf die Untersuchung der endlichen algebraischen Erweiterungen von P zurückgeführt. Und wir bemerken noch, daß jede endliche algebraische Erweiterung von P als Kongruenzkörper K aus einem Körper K erzeugt werden kann.

Besitzt K keine in bezug auf P primitive Größe, so sei es möglich, K aus P durch Adjunktion von $\bar{n}$ Größen zu erhalten, und es sei $\bar{n}_i > 1$ die kleinste derartige Zahl. Nach den Steinitzschen Resultaten ist dann der größte Grad b_i, den eine Größe von K in bezug auf P haben kann, durch p teilbar, und die zugehörige Gleichung, der eine solche Größe genügt, ist rational in z^p. Dann ist $F_i(z)$ einer rationalen Funktion von z^p modulo p kongruent. Ist allgemeiner $F_i(z)$ modulo p einer in $z^{p^{\bar{f}_i}}$ rationalen Funktion kongruent, dagegen keiner, die in $z^{p^{\bar{f}_i+1}}$ rational wäre, so nennen wir $\bar{f}_i$ den Exponent von $\mathfrak{p}_i$. Man kann nun aus den Steinitzschen Sätzen leicht folgern, daß $\frac{g_i}{b_i}$ eine Potenz p^{f_i} von p, und $f_i \leqq (\bar{n}_i - 1) f_i$ ist. Wir nennen f_i den Defekt von $\mathfrak{p}_i$. Andererseits kann man beweisen, daß $\bar{n}_i \leqq n$ ist, wo n die Anzahl der Variabeln x_i ist. Daher kann für $n = 1$ die Irregularität erster Art nicht vorkommen. Endlich folgt aus $f_i \geqq 1$ die Abschätzung $p \leqq \sqrt{a_i b_i} \leqq \sqrt{m}$.

Wir wollen endlich noch den Satz anführen, daß wenn eine Primzahl p in K einen Primidealteiler $\mathfrak{p}_i$ besitzt, der irregulär

---

1) Vgl. z. B. Weber, Lehrbuch der Algebra, Bd. II, (2. Auflage) S. 605—611, Hilbert, Bericht ü. d. Th. d. a. Z., S. 189—191.

2) Steinitz, Algebraische Theorie der Körper, Journal f. d. r. u. a. Math., Bd. 137, S. 167—309.

von der ersten Art ist, dasselbe auch in jeder Erweiterung von K stattfindet. Er ist insofern von Interesse, als er zeigt, daß die Irregularität erster Art in Galoischen Körpern nicht seltener auftritt, als in den allgemeinen.

Da für den Fall $n = 1$, wo also nur eine Variable x_1 in R vorkommt, die Irregularität erster Art nicht auftritt, kann aus der Zerlegung der linken Seite der Fundamentalgleichung modulo p in irreduzible Faktoren auf die Zerlegung der Primzahl p in Primidealteiler und namentlich auf die Werte der Zahlen g_i, e_i geschlossen werden. Dagegen kann auch für $n = 1$ die Fundamentalgleichung zur Untersuchung der Diskriminanten nicht immer angewandt werden, da die Irregularität zweiter Art für jeden positiven Wert von n auftreten kann[1]).

Um ein Beispiel für das Auftreten einer solchen Irregularität zu geben, adjungieren wir zum Körper $R = [x_1]$ die beiden Größen $\xi_1 = \sqrt{x_1}$, $\xi_2 = \sqrt{2}$. Im so entstehenden — normalen — Körper K bilden die Größen

$$1, \xi_1, \xi_2, \xi_1 \xi_2$$

eine Fundamentalbasis für die ganzen Größen, wie man durch einen analogen Ansatz beweisen kann, wie bei dem oben behandelten Beispiel. Die Fundamentalgleichung, der die Fundamentalform

$$v_1 + v_2 \sqrt{x_1} + v_3 \sqrt{2} + v_4 \sqrt{2 x_1}$$

genügt, ist modulo 2 kongruent

$$(z^2 - v_1^2 - v_2^2 x_1 - 2 v_3^2 - 2 v_4^2 x_1)^2.$$

Setzen wir in den irreduziblen Teiler $z^2 - v_1^2 - v_2^2 x_1 - 2 v_3^2 - 2 v_4^2 x_1$ der Fundamentalgleichung für z die Fundamentalform ein, so ist der entstehende Ausdruck gleich

$$2 (v_1 v_2 \sqrt{x_1} + v_1 v_3 \sqrt{2} + v_1 v_4 \sqrt{2 x_1} + \cdots)$$

also durch 2 teilbar. Der größte gemeinsame Teiler von diesem Ausdruck und von 2 ist also gleich $2 = (\sqrt{2})^2$. Daher ist der Primidealteiler von 2 irregulär von der zweiten Art. (Er ist gleich $\sqrt{2}$, was wir aber nicht brauchen.)

---

1) Der zweite Teil des Hilfssatzes, auf dem die Behandlung der irregulären Primzahlen im Buche von König beruht (S. 544—546), besagt in unserer Ausdrucksweise, daß die Irregularität zweiter Art nicht vorkommen kann. Der Fehlschluß besteht darin, daß König aus der Kongruenz $F_i'(\mathfrak{g}) \equiv 0$ (mod. $\mathfrak{p}$) (S. 545, Ende) schließt, daß $F_i(z)$ reduzibel sein müßte. $F_i'(z)$ konnte aber auch identisch verschwinden, falls $F_i(z)$ modulo p einer in z^p rationalen Funktion kongruent ist.

Ist überhaupt ein Primidealteiler $\mathfrak{p}_i$ von p irregulär von der zweiten Art, so ist offenbar $e_i \geqq 2$. Um den größten gemeinsamen Teiler $\mathfrak{p}_i^{\bar{e}_i}$ von p und $F_i(\mathrm{M})$ zu bestimmen, kann man den folgenden Weg einschlagen. Man bilde, ganz analog, wie bei der Untersuchung der Irregularität erster Art die zu $\mathfrak{p}_i$ gehörenden Kongruenzkörper K und P. Es sei nun $\omega_1, \omega_2, \ldots, \omega_{g_i}$ eine Basis von K in bezug auf P, $\omega_1, \bar{\omega}_2, \ldots, \bar{\omega}_{g_i}$ die zu den entsprechenden Kongruenzklassen gehörenden Größen von K, von denen man annehmen kann, daß ihre Nenner zu p teilerfremd sind. Bildet man mit Unbestimmten u_i die Linearform $u_1\bar{\omega}_1 + \cdots + u_{g_i}\omega_{g_i} = \mathrm{T}$ und die modulo p irreduzible Kongruenz $\Pi(z) \equiv 0$ niedrigsten Grades, der T modulo $\mathfrak{p}_i$ genügt, so ist, wie sich unschwer beweisen läßt, der größte gemeinsame Teiler von $\Pi(\mathrm{T})$ und p durch $\mathfrak{p}_i^{\bar{e}_i}$, nicht aber durch $\mathfrak{p}_i^{\bar{e}_i+1}$ teilbar. — Wir bemerken noch, daß stets $a_i \geqq e_i$, andererseits $F_i(z)$ modulo p einer in z^p rationalen Funktion kongruent ist, falls $\mathfrak{p}_i$ irregulär von der zweiten Art ist. Daher ergibt sich für p die Abschätzung $p \leqq b_i \leqq \frac{n}{2}$.

Die Zahl d^2 ist Norm eines Idealteilers und ist eine wichtige Invariante des Körpers. Sie hängt — entgegen einer Bemerkung von J. König[1]) — von der Wahl der Fundamentalbasis nicht ab. Um ihre Zusammensetzung zu untersuchen, kann man die Methoden von Dedekind und Hilbert zur Untersuchung der Ringe in folgender Weise übertragen.

Ist M eine — festgewählte — Fundamentalform des Körpers, so betrachte man die Gesamtheit aller ganzen Größen von K, die sich als Polynome in M so ausdrücken lassen, daß die Koeffizienten rationale Funktionen in den x_i und v_i sind, deren Generalnenner ein primitives Polynom der v_i[2]) ist. Diese Gesamtheit nennen wir einen Fundamentalring $\mathfrak{R}$ von K. Man stelle ferner alle ganzen Größen von K auf, die die Eigenschaft haben, daß jede ganze Größe von K nach Multiplikation mit irgend einer von ihnen zu $\mathfrak{R}$ gehört. Den größten gemeinsamen Idealteiler dieser Größen nennen wir den Führer φ von $\mathfrak{R}$. Er ist eine Invariante des Körpers und seine Norm ist gleich d^2. Der Führer φ kann auch

---

1) König, S. 548 Mitte.

2) Dies bedeutet, daß der größte gemeinsame Teiler aller ganzzahligen Polynome der x_i, die als Koeffizienten der Potenzprodukte der v_i auftreten, gleich 1 ist.

dadurch definiert werden, daß das Produkt von φ mit der Differente des Körpers gleich dem größten gemeinsamen Idealteiler der Koeffizienten aller Potenzprodukte der Unbestimmten v_i in $\Phi'(\mathrm{M})$ ist. — Das obige Produkt ließe sich auch als ein Produkt der Elemente des Körpers[1]) definieren. —

Die Bestimmung der höchsten Potenz einer irregulären Primzahl p, die in d^2 auftritt, hängt aufs engste zusammen mit der Zerlegung der linken Seite der Fundamentalgleichung in irreduzible Faktoren nach den höheren Potenzen von p als Modul. Wir knüpfen an die bereits oben benutzten Begriffsbildungen an: Wir bezeichnen also eine Größe von R als ganz in bezug auf p, wenn sie sich als ein Quotient von ganzen Größen von R darstellen läßt, sodaß der Nenner zu p teilerfremd ist. Von hier aus können wir die Teilbarkeit in bezug auf p definieren, und ebenso unter den Größen von K ganze in bezug auf p auszeichnen. Die Fundamentalbasis, Fundamentalform, Fundamentalgleichung und Fundamentaldiskriminante, ebenso die Körperdiskriminante lassen sich dann in bezug auf p ganz analog definieren. In der so entwickelten Idealtheorie kommen aber nur Primidealteiler von p vor, da alle übrigen Primgrößen von R in bezug auf p als Einheiten aufzufassen sind[2]). Und nun läßt sich beweisen, daß diese Primidealteiler von p eineindeutig den Primidealteilern von p in der Kroneckerschen Theorie entsprechen, so daß die Zahlen g_i, e_i dieselben sind, ebenso sind die Zahlen a_i, b_i dieselben. Ebenso sind aber die genauen Potenzen von p, die in der Fundamentaldiskriminante, bzw. Körperdiskriminante in der Kroneckerschen Theorie vorkommen, gleich den Potenzen von p, die in der Fundamentaldiskriminante, bzw. Körperdiskriminante in bezug auf p vorkommen, so daß wir uns auf die Untersuchung dieser Größen in bezug auf p beschränken können[3]). Ist M die Fundamentalform von K, so kann man die Fundamentalbasis in bezug auf p auch mit Hilfe von Potenzen von M bilden, unter Adjunktion der Unbestimmten v_i, und zwar aus den m Ausdrücken

---

1) Vgl. Hilbert, Bericht ü. d. Th. d. a. Z., S. 200.

2) Diese Begriffsbildungen sind bekannt und vor allem für den Fall von Zahlkörpern wiederholt benutzt worden. Zwischen solchen Kongruenzbestimmungen für den Fall der Zahlkörper, und für den Fall der Funktionenkörper besteht aber der fundamentale Unterschied, daß im zweiten Falle eine Kongruenzklasse modulo p im allgemeinen keine einzige im gewöhnlichen Sinne ganze Größe enthält.

3) Mit Hilfe dieser Auffassungen ist es möglich, die gewöhnlichen Methoden zur Begründung der Idealtheorie auf alle in Betracht kommenden Fälle auszudehnen, z. B. auch im Falle unendlicher Körper.

$$(6)\quad \tau_0 = 1,\ \tau_i = \frac{a_i + a_i' \mathrm{M} + \cdots + a_i^{(i-1)} \mathrm{M}^{i-1} + \mathrm{M}^i}{p^{f_i}} = \frac{\psi_i(\mathrm{M})}{p^{f_i}} \qquad i = 1, \ldots, m-1.$$

Dann kommt p in d genau zur Potenz $f_1 + f_2 + \cdots + f_{m-1}$ vor. Die Zahl p^{f_i} kann man aber auch direkt definieren, als die höchste Potenz von p, die $\psi_i(\mathrm{M})$ teilen kann, wenn $\psi_i(z)$ ein Polynom i-ten Grades in z ist mit in bezug auf p ganzen Koeffizienten aus R und dem höchsten Gliede z^i. Solche Polynome $\psi_i(z)$, die durch die höchst mögliche Potenz von p teilbar sind, kann man aus den Teilern der linken Seite der Fundamentalgleichung $\Phi(z)$ nach höheren Potenzen von p als Modul zusammensetzen. Hierbei ist vor allem der Satz zu benutzen: Für jedes ganze positive k gilt eine Kongruenz

$$\Phi(z) \equiv F_1^{(k)}(z) \ldots F_i^{(k)}(z) \ldots F_s^{(k)}(z) \qquad (\mathrm{mod.}\ p^k),$$

in der $F_i^{(k)}(z)$ ein Polynom in z vom Grade $a_i b_i = g_i e_i$ ist, mit in bezug auf p ganzen Koeffizienten aus $R(v_i)$ und dem höchsten Koeffizienten 1. Dabei kann angenommen werden, daß die Kongruenzen gelten:

$$F_i^{(k)}(z) \equiv F_i^{(k')}(z)\ (\mathrm{mod.}\ p^{k'}) \quad (k' < k), \qquad F_i^{(k)}(z) \equiv F_i^{(1)}(z) \equiv F_i^{a_i}(z) \quad (\mathrm{mod.}\ p).$$

Hieraus folgt dann ohne Schwierigkeit:

$$F_i^{(k)}(\mathrm{M}) \equiv 0 \ \left(\mathrm{mod.}\ \mathfrak{p}_i^{e_i k}\right), \qquad F_i^{(k)}(\mathrm{M}) \not\equiv 0 \ (\mathrm{mod.}\ \mathfrak{p}_j) \qquad (j \neq i).$$

Die Fundamentalform M genügt also nach jeder Potenz von $\mathfrak{p}_i$ als Modul einer Kongruenz vom Grade $a_i b_i$ — und keiner Kongruenz vom niedrigeren Grade mit in bezug auf p ganzen Koeffizienten und dem höchsten Koeffizienten 1 nach einer hinreichend großen Potenz von $\mathfrak{p}_i$ als Modul (z. B. von der höchsten in d als Teiler vorkommenden Potenz von $\mathfrak{p}_i$ an).

Wir stellen nun für jedes $j \leqq a_i$ die größte Zahl m_i^j von der Eigenschaft auf, daß es ein Polynom $W_i^{(j)}(z)$ vom Grade jb_i mit in bezug auf p ganzen Koeffizienten aus $R(v_i)$ und dem höchsten Koeffizienten 1 gibt, welches modulo p kongruent $F_i^j(z)$ ist, und für welches $W_i^{(j)}(\mathrm{M})$ durch $\mathfrak{p}_i^{m_i^{(j)}}$ teilbar ist. $m_i^{(a_i)}$ kann man gleich ∞ setzen. Diese Zahlen $m_i^{(j)}$ sind von der Wahl der Fundamentalbasis unabhängige Invarianten des Körpers. Durch sie lassen sich alle für uns in betracht kommenden Größen ausdrücken, insbesondere die Zahlen f_i in (6) und der Führer φ des Fundamentalringes. Betrachten wir ein in (6) vorkommendes Polynom

$\psi_t(z)$, für welches $f_t > 0$ ist, so ist es nach einer beliebig hohen Potenz von p einem Produkt von $s+1$ Faktoren kongruent. Die ersten s-Faktoren sind modulo p kongruent gewissen Potenzen der $F_i(z)$ und können direkt als gewisse Polynome $W_i^{(j_{t,i})}$ angenommen werden, wenn man $\psi_t(z)$ in geeigneter Weise abändert. Der $(s+1)$-te Faktor ist modulo p zu allen $F_i(z)$ teilerfremd und bis auf seinen Grad willkürlich. Für die $j_{t,i}$ gelten die Ungleichungen:

$$j_{t,i} \leqq a_i,\ f_t \leqq \frac{m_i^{(j_{t,i})}}{e_i} \quad (i = 1, 2, \dots s), \quad j_{t,1} b_1 + j_{t,2} b_2 + \dots + j_{t,s} b_s \leqq t,$$

mit deren Hilfe der Wert von f_t berechnet werden kann als die größte ganze Zahl, die diesen Bedingungen genügt. Insbesondere erhält man als Bedingung für $f_t \geqq 1$, d. h. für die Existenz eines durch p teilbaren Polynoms in M vom Grade t und mit dem höchsten Koeffizient 1 die Relationen

$$m_i^{(j_{t,i})} \geqq e_i \qquad (i = 1, 2, \dots, s).$$

Andererseits läßt sich der Führer des Fundamentalringes direkt bestimmen. Sein Wert ist, wie eine einfache Überlegung zeigt:

$$\prod_{i,p} \mathfrak{p}_i^{\,m_i^{(a_i-1)} - e_i + 1},$$

wo das Produkt über alle irregulären Primzahlen p und über alle Indizes i, denen ein irregulärer Primteiler entspricht, zu erstrecken ist. Als die Norm hiervon erhält man:

$$d^2 = \prod_p \frac{1}{p^m} \prod_i p^{g_i m_i^{(a_i-1)} + g_i}.$$

Da für die m_i die folgenden Relationen gelten, solange die oberen Indizes a_i nicht übersteigen:

$$m_i^{(j+j')} \geqq m_i^{(j)} + m_i^{(j')}, \quad m_i^{(jj')} \geqq j m_i^{(j')}, \quad m_i^{(jj')} \geqq j' m_i^{(j)},$$

erhält man für die im Führer vorkommende Potenz von $\mathfrak{p}_i$ die Abschätzung

$$\geqq (a_i - 1)\, \bar{e}_i - (e_i - 1),$$

wo $\mathfrak{p}_i^{\bar{e}_i}$ der größte gemeinsame Teiler von p und $F_i(\mathrm{M})$ ist. — Die im Obigen entwickelten Methoden können auch zur Untersuchung der Zusammensetzung der Körperdiskriminante angewandt werden, worauf hier jedoch nicht weiter eingegangen werden soll. —

Wir wollen nun zeigen, wie die Hilbertschen Diskriminantensätze für unseren allgemeineren Fall zu beweisen sind. Es sei K' ein Körper über K vom Grade m'. Es seien $\Delta, \Delta', \Delta''$ die Diskriminanten von K, von K' und die Relativdiskriminante von K' in bezug auf K, j, j', j'' die Differenten von K, K' und die Relativdifferente von K' in bezug auf K, endlich mögen die Quotienten $\frac{\Delta}{j}, \frac{\Delta'}{j'}, \frac{\Delta''}{j''}$ durch S, S', S'' bezeichnet werden. Dann wird es sich um den Beweis der Gleichungen

$$\Delta' = \Delta^{m'} N\Delta'', \quad j' = jj''$$

handeln, wo durch N die Norm in K in bezug auf R bezeichnet wird. Ist p eine Primzahl, so bezeichnen wir die in $\Delta, \Delta', \Delta'', j, j', j''$, S, S', S'' aus den Teilern von p bestehenden Bestandteile durch $\Delta_p, \Delta'_p, \Delta''_p, j_p, j'_p, j''_p, S_p, S'_p, S''_p$. Es genügt dann zu beweisen:

$$\Delta'_p = \Delta_p^{m'} N\Delta''_p, \quad j'_p = j_p j''_p \text{ [1]}.$$

Nun besteht eine Fundamentalbasis für die in bezug auf p ganzen Größen von K aus m Größen $\omega_1, \ldots, \omega_m$. Andererseits kann man durch sukzessive Reduktion zeigen, daß K' in bezug auf K eine aus m' Elementen bestehende relative Fundamentalbasis für die in bezug auf p ganzen Größen besitzt:

$$\omega''_1, \omega''_2, \ldots, \omega''_{m'},$$

sodaß jede in bezug auf p ganze Größe von K' sich in der Form

$$\alpha_1 \omega''_1 + \cdots + \alpha_{m'} \omega''_{m'}$$

darstellen läßt, wo $\alpha_1, \ldots, \alpha_{m'}$ in bezug auf p ganze Größen von K sind. Dann besteht, wie hieraus unmittelbar folgt, die Fundamentalbasis von K' für die in bezug auf p ganzen Größen aus den mm' Elementen

$$\omega_i \omega''_k \qquad i = 1, \ldots, m; \quad k = 1, \ldots, m'.$$

Wir betrachten nun die Determinanten $\Omega = |\omega_i|$, $\Omega'' = |\omega''_k|$, $\Omega' = |\omega_i \omega''_k|$, deren Zeilen aus der ersten durch den Übergang zu den konjugierten Größen hervorgehen. Ersetzen wir die Elemente der ersten Zeile in $\Omega, \Omega'', \Omega'$ bzw. durch $x_1, \ldots, x_m$; $y_1, \ldots, y_{m'}$; $x_1 y_1, \ldots x_1 y_{m'}, x_2 y_1, \ldots x_m y_{m'}$, so mögen die entstehenden Deter-

[1] Für die Primteiler der Diskriminante, die keine Zahlen sind, folgen die entsprechenden Gleichungen bereits aus der Kroneckerschen Theorie der Fundamentalgleichung.

minanten durch $\Omega(x_i)$, $\Omega''(y_k)$, $\Omega'(x_i y_k)$ bezeichnet werden. Und nun gelten die leicht zu beweisenden Sätze: Ω^2, Ω''^2, Ω'^2 sind genau durch Δ_p, Δ''_p, Δ'_p teilbar, die größten gemeinsamen Teiler aller Koeffizienten in $\Omega\Omega(x_i)$, $\Omega''\Omega''(y_k)$, $\Omega'\Omega'(x_i y_k)$ sind bzw. genau durch S_p, S'_p, S''_p teilbar. Andererseits gilt die Gleichung

$$\Omega'(x_i, y_k) = \Omega^{m'-1}\Omega(x_i)\frac{N(\Omega'')}{\Omega''}\Omega''(y_k),$$

wo $N\Omega''$ das Produkt von Ω'' mit den entsprechenden Ausdrücken bedeutet, die für die in bezug auf R zu K konjugierte Körper gebildet sind. Diese Gleichung geht, wenn man die ω_i, ω''_k und alle konjugierten Größen als Unbestimmte auffaßt, in eine Determinantenidentität über, die durch bekannte Schlüsse sofort zu beweisen ist. Aus ihr folgt, wenn wir x_i durch die ω_i, y_k durch die ω''_k ersetzen,

(7) $$\Omega' = \Omega^{m'} N\Omega'', \quad \Delta'_p = \Delta_p^{m'} N\Delta''_p.$$

Daher erhalten wir die weitere Gleichung

$$\Omega'\Omega'(x_i, y_k) = \Omega^{2(m'-1)}\Omega\Omega(x_i)\frac{N(\Omega''^2)}{\Omega''^2}\Omega''\Omega''(y_k),$$

aus der durch Vergleichung der größten gemeinsamen Teiler der Koeffizienten der Unbestimmten auf beiden Seiten folgt

$$S'_p = \Delta^{m'-1} S_p \,.\, \frac{N\Delta''_p}{\Delta''_p} S''_p.$$

Und hieraus und aus (7) ergibt sich endlich

$$\mathfrak{j}'_p = \mathfrak{j}_p \mathfrak{j}''_p.$$

Von hier aus kann nun die Theorie der zu den einzelnen Primidealteilern in normalen Körpern gehörenden besonderen Körper (Trägheits-, Zerlegungs- und Verzweigungskörper) entwickelt werden. Wir gehen darauf nicht weiter ein, bemerken aber, daß diese Theorie gegenüber der entsprechenden im Falle der Zahlkörper bedeutende Unterschiede aufweist, namentlich für Primidealteiler $\mathfrak{p}_i$, die die Eigenschaft haben, daß die zugehörigen Faktoren $F_i(z)$ der Fundamentalgleichung modulo p einer in z^p rationalen Funktion kongruent sind (für die daher p jedenfalls auch ein Diskriminantenteiler ist).

In jedem endlichen Körper K können natürlich nur endlich viele Primzahlen irregulär sein. Indessen gibt es in der Regel überhaupt nur endlich viele natürliche Primzahlen, die in K weiter

zerfallen, sobald K Variabeln enthält. Genauer gilt folgender Satz:

Enthält ein endlicher algebraischer Körper K über $R = [x_i]$ keine nicht rationalen algebraischen Zahlen, so zerfallen nur endlich viele natürliche Primzahlen in K weiter[1]). Ist allgemeiner $\mathfrak{K}$ der größte in K enthaltene algebraische Zahlkörper, so gibt es nur endlich viele Primideale in $\mathfrak{K}$, die in K weiter zerfallen.

Es sei α eine ganze primitive Größe von K in bezug auf den Körper $(\mathfrak{K}, x_i)$, und sie genüge der in $(\mathfrak{K}, x_i)$ irreduziblen Gleichung

$$F(z) = 0$$

vom Grade μ und mit dem höchsten Koeffizient 1, deren übrige Koeffizienten ganze Größen von $(\mathfrak{K}, x_i)$ sind. Ist nun ein Primidealteiler $\mathfrak{p}$ von $\mathfrak{K}$ (und daher auch von $(\mathfrak{K}, x_i)$) in K nicht unzerlegbar, sondern zerfällt weiter, so ist das Polynom F in den Variabeln x_i, z modulo $\mathfrak{p}$ reduzibel und zerfällt in mehrere Faktoren, die Polynome in x_i, z sind, und deren Anzahl wenigstens gleich ist der Anzahl der zu K gehörenden Primidealteiler von $\mathfrak{p}$. Und es ist also nur zu beweisen, daß es nur endlich viele Primidealteiler in $\mathfrak{K}$ gibt, nach denen als Moduln F zerfällt. Nun enthält der Körper $(\mathfrak{K}, x_i, \alpha)$ keine weiteren Zahlenirrationalitäten, die nicht in $\mathfrak{K}$ enthalten wären. Daraus läßt sich zeigen, daß das Polynom F in x_i, z absolut irreduzibel ist[2]).

Wir brauchen also nur den folgenden Hilfssatz zu beweisen: Ist $F(x_i, z)$ ein absolut irreduzibles Polynom mit Koeffizienten aus einem Zahlkörper $\mathfrak{K}$, so gibt es in $\mathfrak{K}$ nur endlich viele Primidealteiler, nach denen F nicht irreduzibel bleibt.

Zum Beweise bemerken wir, daß die Frage, ob $F(x_i, z)$ einen Teiler mit vorgegebenen genauen Gradzahlen in x_i, z besitzt, sich darauf zurückführen läßt, ob ein gewisses Polynom

$$T(t, y_i) = t^m + Y_1(y_i) t^{m-1} + \cdots + Y_m(y_i), \tag{8}$$

wo $Y_k(y_i)$ Polynome in gewissen Variabeln y_i mit rationalen Koeffizienten aus $\mathfrak{K}$ sind, einen in bezug auf t linearen Teiler hat. Und

---

1) Diesen Satz habe ich bereits im Winter 1917—18 gefunden und am 18. Februar 1918 Herrn E. Landau brieflich mitgeteilt.

2) Den einfachen Beweis dieser Tatsache werde ich an einer anderen Stelle veröffentlichen, und bei dieser Gelegenheit auch auf viele merkwürdigen Folgerungen eingehen, die sich an sie knüpfen lassen.

zwar hat $F(x_i, z)$ dann und nur dann einen Teiler mit jenen Gradzahlen, und mit Koeffizienten aus einem Zahlkörper K, wenn T in K einen in t linearen, in den y_i ganzen Teiler hat. Diese von F. Mertens durchgeführte Zurückführung[1]) gilt nun auch für die Teilbarkeit modulo eines Primidealteilers von $\mathfrak{K}$, wenn dieser Primidealteiler nur keine zu kleine Norm hat. Es ergibt sich insbesondere: $F(x_i, z)$ besitzt modulo eines Primidealteilers $\mathfrak{p}$ von $\mathfrak{K}$ mit hinreichend großer Norm dann und nur dann einen Teiler mit jenen genauen Gradzahlen, wenn $T(t, y_i)$ modulo $\mathfrak{p}$ einen in t linearen Teiler besitzt. Wir brauchen also nur zu beweisen, daß jedes Polynom von der Form (8), welches im Bereiche aller Zahlen keinen in t linearen Teiler besitzt, auch nur nach endlich vielen Primidealteilern von $\mathfrak{K}$ als Modul einen in t linearen Teiler besitzen kann.

Wir benutzen nun den folgenden Hilfssatz: Es seien $\Phi_1(t, y_i)$, $\Phi_2(t, y_i)$, ... $\Phi_r(t, y_i)$ absolut irreduzible Polynome in t und y_i. Dann kann man, nach einer geeigneten ganzzahligen linearen Substitution von der Determinante 1 der y_i, allen y_i bis auf y_1 solche ganzzahlige Werte beilegen, daß die entstehenden Polynome $\overline{\Phi}_1(t, y_1)$, ..., $\overline{\Phi}_r(t, y_1)$ absolut irreduzibel bleiben[2]).

Wenden wir diesen Satz auf alle absolut irreduziblen Teiler von T an, so entsteht aus $T(t, y_i)$ ein Polynom $\mathrm{T}(t, y_1)$, welches ebenfalls keinen in t linearen Teiler besitzt. Und es ist dann zu beweisen, daß es nur endlich viel Primidealteiler in $\mathfrak{K}$ gibt, nach denen $\mathrm{T}(t, y_1)$ einen in t linearen Teiler enthält. Es sei $f(y_1) = c(y_1 - \alpha_1) \dots (y_1 - \alpha_k)$ das von t freie Glied in T; wir adjungieren zu $\mathfrak{K}$ alle Wurzeln $\alpha_1, \dots, \alpha_k$ von f, und beweisen, was offenbar genügt, unseren Satz für den so entstehenden Körper, den wir wieder durch $\mathfrak{K}$ bezeichnen. Hat nun T modulo eines Primidealteilers $\mathfrak{p}$ von $\mathfrak{K}$ eine Wurzel $t - \varphi(y_1)$, so muß $\varphi(y_1)$ modulo $\mathfrak{p}$ ein Teiler von $f(y_1)$ und daher einem Ausdruck

$$\gamma(y_1 - \alpha') \, y_1 - \alpha'') \dots (y_1 - \alpha^{(i)}) \tag{9}$$

kongruent sein, in dem $(y_1 - \alpha') \dots (y_1 - \alpha^{(i)})$ ein Teiler von $f(y_1)$ ist. Um γ genauer festzulegen, setzen wir in T für t an: $\gamma y_1^i + \cdots$ und vergleichen die Koeffizienten der höchsten vorkommenden

1) Vgl. F. Mertens, Wiener Ber., Bd. 120, S. 1485 ff.

2) Den Beweis dieses Satzes werde ich in einer Abhandlung: Über absolute Irreduzibilität und konvexe Polyeder darstellen, die an einer anderen Stelle erscheinen wird.

Potenz von y_1. So kommen wir zu einem Ausdruck $P_i(\gamma)$, und der Koeffizient von y_1^i in $\varphi(y_1)$ ist eine Wurzel der Kongruenz $P_i(\gamma) \equiv 0$ (mod. $\mathfrak{p}$). Nun aber bemerken wir, daß die Koeffizienten der Funktion $P_i(\gamma)$ nur von i, nicht aber von $\mathfrak{p}$ abhängen. Adjungieren wir dem Körper $\mathfrak{K}$ alle Wurzeln der Gleichungen

$$(10) \qquad P_0(z) = 0, \quad P_1(z) = 0, \ldots, P_k(z) = 0,$$

so genügt es, unseren Hilfssatz für den so entstehenden Körper zu beweisen, den wir wiederum mit $\mathfrak{K}$ bezeichnen. Dann muß aber der höchste Koeffizient von y_1 in $\varphi(y_1)$ einer der endlich vielen Wurzeln der Polynome (10) kongruent sein. Wir erhalten daher aus (9) endlich viele Funktionen

$$\psi_1(y_1), \psi_2(y_1), \ldots, \psi_s(y_1)$$

von der Eigenschaft, daß jedesmal, wenn T für irgend einen Primidealteiler $\mathfrak{p}$ von $\mathfrak{K}$ einen in t linearen Teiler $t - \varphi(y_1)$ bekommt, $\varphi(y_1)$ kongruent einem der Polynome $\psi(y_1)$ modulo $\mathfrak{p}$ sein soll. Gibt es unendlich viele solche Primidealteiler $\mathfrak{p}$ in $\mathfrak{K}$, so muß es eines unter den ψ_i geben, etwa $\psi_1(y_1)$, welches für unendlich viele Primidealteiler von $\mathfrak{K}$ der Kongruenz genügt

$$\mathrm{T}(\psi_1(y_1), y_1) \equiv 0.$$

Dann müßte aber die Gleichung gelten $\mathrm{T}(\psi_1(y_1), y_1) = 0$, und $t - \psi_1(y_1)$ wäre Teiler von $\mathrm{T}(t, y_1)$, entgegen der Annahme, w. z. b. w.

Über die akzidentelle Doppelbrechung im Zelloidin und in der Zellulose.

Von

H. Ambronn, Jena.

(Mitteilung aus dem Institut für Mikroskopie an der Universität Jena.)

Vorgelegt von C. Runge in der Sitzung vom 28. März 1919.

Die starken Abweichungen in der Reihenfolge der Interferenzfarben, wie sie die Apophyllite von verschiedenen Fundorten zeigen, sind den Kristallographen seit langem bekannt. Es gibt Kristalle dieses Minerals, die für alle Farben positive und andere, die negative Doppelbrechung besitzen; und zwischen diesen beiden Grenzfällen sind auch alle Übergänge bekannt. Schon H. de Sénarmont[1]) hat diese auffallenden Erscheinungen nachzuahmen versucht, indem er Mischkristalle aus dem positiven Bleidithionat und dem negativen Strontiumdithionat herstellte. Auch an diesen lassen sich zwischen den beiden Grenzfällen, die dem Verhalten der reinen Salze entsprechen, alle Übergänge beobachten. Hier sind die optischen Eigenschaften der beiden Komponenten genau bekannt, und aus dem Zusammenwirken in den Mischkristallen lassen sich die Aufeinanderfolgen der sog. Brucit-, Leukozyklit-, Andreasberger- und Chromozyklitfarben leicht übersehen. Bei den Apophylliten kennt man dagegen die beiden Komponenten

---

Bemerkung: Die vorliegende Mitteilung gibt in kurzer Zusammenfassung den Inhalt von zwei Vorträgen wieder, die ich auf Einladung der Kgl. Ges. d. Wiss. im Januar 1918 im Hörsaal des physikalischen Instituts der Universität Göttingen gehalten habe.

1) H. de Sénarmont, Pogg. Ann. 86, 35, 1852.

nicht in reinem Zustande, aber schon C. Klein[1]) schloß mit Recht aus dem optischen Verhalten, daß die Apophyllite ebenfalls als Mischkristalle aus einer positiven und einer negativen Komponente zu betrachten seien. Neuere Untersuchungen[2]) haben auch quantitativ die Zusammensetzung der einzelnen Farbenfolgen aufgeklärt, wenn auch die Frage nach der Art der beiden Komponenten bis jetzt noch offen geblieben ist.

Dieselben Änderungen in der Reihenfolge der Interferenzfarben treten aber auch auf, wenn keine Mischkristalle vorliegen, die aus zwei an sich schon doppelbrechenden Komponenten von entgegengesetztem Vorzeichen der Doppelbrechung bestehen. Man kann sich den Fall denken, daß zwei Körper mit einander gemischt werden, die im reinen Zustande bei Spannungen akzidentelle Doppelbrechung von entgegengesetztem Charakter ergeben. Es wird dann quantitativ auf die Stärke der beiden Anisotropieen und auf das Mischungsverhältnis ankommen, wie sich das Zusammenwirken gestaltet. Genauere Untersuchungen über solche Fälle liegen zwar noch nicht vor, doch möchte ich einige Beispiele anführen, in denen derartige Verhältnisse jedenfalls in Betracht kommen.

Es ist seit langem bekannt, daß Würfel von Chlornatrium durch Druck in der Richtung der Würfelnormalen negativ doppelbrechend werden[3]), während Würfel aus Chlorkalium das entgegengesetzte Verhalten zeigen. Man kann nun aus beiden Körpern Mischkristalle herstellen — es kommen solche auch in der Natur vor —, die zwischen den beiden Grenzfällen Übergänge zeigen.

Zweitens sind von F. Pockels[4]) einige stark bleihaltige Gläser untersucht worden, die sich ähnlich wie Chlorkaliumwürfel verhalten, d. h. bei Druck positiv doppelbrechend werden, während schwächer bleihaltige Gläser das entgegengesetzte Verhalten zeigen. Das Vorzeichen der akzidentellen Doppelbrechung solcher Gläser hängt von der Höhe des Bleigehaltes ab. Bei einem bestimmten Bleigehalt bleiben sie wenigstens für eine Wellenlänge isotrop; für die anderen Wellenlängen besteht noch schwache Anisotropie. Es müssen also auch hier starke Abweichungen in der Dispersion der Doppelbrechung auftreten, was Pockels auch schon bestätigen konnte.

1) C. Klein, Mineralog. Mitteil. XIII, N. Jahrb. f. Miner. 2, 176, 1892.
2) A. Wenzel, N. Jahrb. f. Miner. Beilage Bd. XLI, 565—627, 1917.
3) R. Brauns, N. Jahrb. f. Miner. 1, 233, 1886.
4) F. Pockels, Ann. d. Phys. (4) 7, 745, 1902.

Ferner möchte ich noch die Mischungen aus arabischem Gummi und Kirschgummi als Beispiel anführen[1]). Zieht man aus zähflüssigem arabischen Gummi Fäden, so erweisen sie sich als optisch positiv, während in gleicher Weise hergestellte Fäden aus Kirschgummi negativ sind. Durch Mischungen erhält man auch hier alle Übergänge, sodaß bei einem bestimmten Mischungsverhältnis die Fäden überhaupt isotrop bleiben.

Hierher gehören auch die eigentümlichen Erscheinungen in der akzidentellen Doppelbrechung des mit echtem Kampfer hergestellten Zelluloids[2]). Streifen aus diesem Körper, die zunächst optisch isotrop sind, zeigen bei Dehnung innerhalb der Elastizitätsgrenze die normale positive Doppelbrechung; tritt jedoch durch stärkere Spannung bleibende Deformation ein, so kehrt sich das Vorzeichen der Doppelbrechung um. Löst man jetzt den Kampfer heraus, so werden sie wieder positiv; läßt man diese entkampferten Streifen längere Zeit in einer gesättigten Kampferlösung liegen, in der sich noch Kampfer im Überschuß befindet, so werden sie von neuem negativ. Es scheint also wieder fester Kampfer eingelagert zu werden. Gerade beim Zelluloid kann man besonders in dickeren Streifen, sehr gut die starken Veränderungen in der Dispersion der Doppelbrechung verfolgen; es treten auch hier ganz dieselben anomalen Farbenfolgen wie bei den Apophylliten und den Mischkristallen aus Blei- und Strontiumdithionat auf.

1. Akzidentelle Doppelbrechung im Zelloidin.

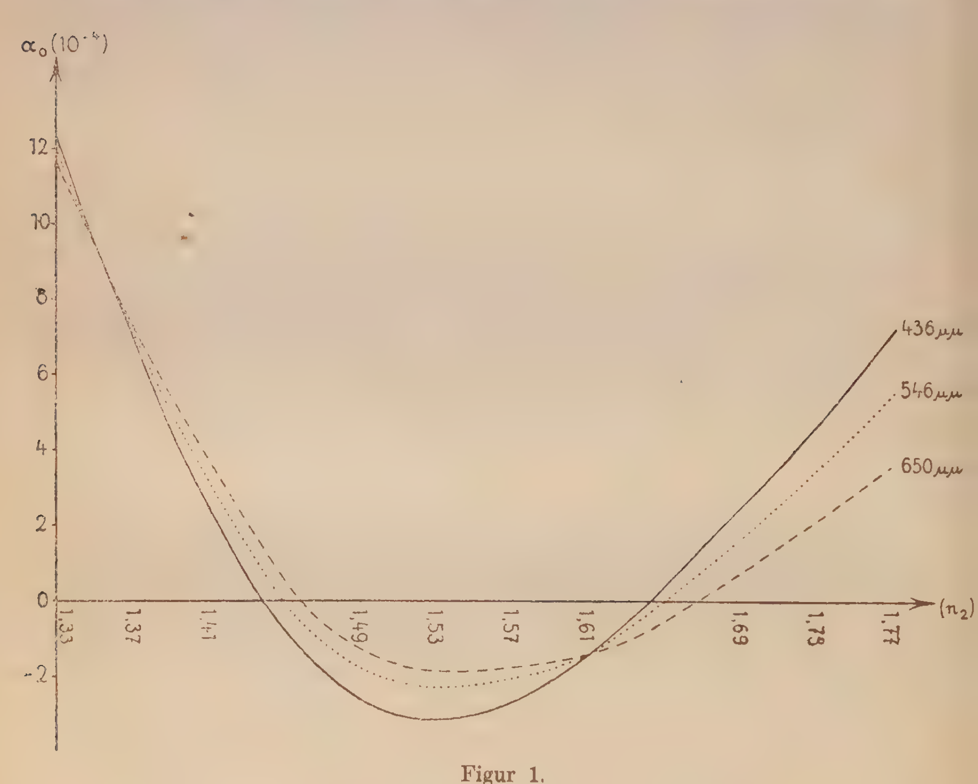
In allen angeführten Beispielen handelt es sich offenbar um Mischungen, in denen die beiden Komponenten entgegengesetztes Vorzeichen der Doppelbrechung besitzen. Wesentlich anders liegen aber die Dinge beim reinen Zelloidin, das aus einer Nitrozellulose von hoher Nitrierungsstufe besteht und mit Flüssigkeit imbibiert ist. Streifen aus diesem Körper zeigen je nach dem Brechungsexponenten der Flüssigkeit, mit der sie durchtränkt sind, ein ganz verschiedenes Verhalten. Sie besitzen eine große Plastizität und lassen sich ohne Schwierigkeit auf über 100 % ihrer ursprünglichen Länge bleibend deformieren. Allerdings liegt auch hier ein System vor, das aus zwei Komponenten aufgebaut ist, und zwar bildet die Imbibitionsflüssigkeit den größeren Anteil — im frischen Zelloidin etwa 75 % —, aber diese Komponente ist stets flüssig,

---

1) H. Ambronn, Wiedem. Ann. **38**, 150, 1889.

2) Ders. Kgl. Sächs. Ges. d. Wiss. math.-phys. Kl. **63**, 249 u. 402, 1911; Kolloidzeitschr. **9**, 147, 1911.

während die andere fest bleibt. Wählt man Flüssigkeiten, die die Nitrozellulose weder chemisch noch physikalisch verändern, so kann man innerhalb der Brechungsexponenten 1,33 und 1,77 alle Übergänge herstellen und erhält dabei das merkwürdige Ergebnis, daß bei diesen beiden Grenzwerten und in ihrer Nähe die akzidentelle Doppelbrechung für alle Farben positiv ist, während in der Nähe des Wertes 1,53 ebenfalls für alle Farben das Vorzeichen umgekehrt wird. Daraus geht schon ohne weiteres hervor, daß für zwei andere Werte — wenigstens jeweils für eine Wellenlänge — die Streifen isotrop sein müssen. Dies ist für Grün etwa bei 1,45 und 1,65 der Fall. Zur Abkürzung möge weiterhin die Stärke der Doppelbrechung $n_a - n_0$ mit α_0 und der Brechungsexponent der Nitrozellulose mit n_1, der der Flüssigkeit mit n_2 bezeichnet werden. Dann ist α_0 für alle Werte von n_2 zwischen 1,33—1,45 und zwischen 1,65—1,77 positiv dagegen für die Werte von 1,45—1,65 negativ, und bei 1,45 und 1,65 selbst ist $\alpha_0 = 0$.



Figur 1.

Die größten absoluten Werte im negativen Gebiete erreicht α_0 in der Nähe von 1,53, wenn nämlich $n_1 - n_2 = 0$ wird. Ob $n_1 - n_2 >$ oder < 0 ist, das ändert an dem Verhalten gar nichts; denn in beiden Fällen kann α_0 sowohl positiv als negativ als auch 0 sein. Einen Überblick über den Verlauf der Werte von α_0 für verschiedene Wellenlängen gibt die beistehende graphische Darstellung (Fig. 1), in der auf der Abszissenachse die Brechungsexponenten der Flüssigkeit für die *D*-Linie und als Ordinaten die Stärken der Doppelbrechung für die Wellenlängen 436 $\mu\mu$, 546 $\mu\mu$, 650 $\mu\mu$ in 10^{-4} eingetragen sind; man ersieht hieraus auch zugleich den anomalen Gang der Dispersion der Doppelbrechung. Wegen anderer Einzelheiten muß ich auf meine früheren Veröffentlichungen verweisen [1]).

Aus diesem ganz eigentümlichen Verhalten, das nach den bisherigen Kenntnissen über den Verlauf der akzidentellen Doppelbrechung kaum verständlich zu machen wäre, glaubte ich den Schluß ziehen zu dürfen, daß in dem System Nitrozellulose + Flüssigkeit ebenfalls ein Zusammenwirken zweier optischer Anisotropieen von entgegengesetztem Vorzeichen stattfinden müsse, zumal auch alle jene Abweichungen in der Reihenfolge der Interferenzfarben auftreten, wie sie z. B. für die Apophyllite und die Mischkristalle aus Blei- und Strontiumdithionat charakteristisch sind [2]).

2. Stäbchendoppelbrechung.

Körper, deren Doppelbrechung vom Brechungsexponenten einer sie durchdringenden Flüssigkeit abhängig ist, sind schon seit langer Zeit bekannt. Die Kieselschalen vieler Diatomeen zeigen, wenn sie in Luft liegen, eine deutliche Doppelbrechung, deren Stärke sinkt, wenn sie in Wasser oder Alkohol gebracht werden, und die ganz verschwindet, wenn sie von einem Medium durchtränkt sind, dessen Brechungsexponent etwa 1,5 ist [3]). Die Doppelbrechung tritt dagegen in Monobromnaphthalin oder Realgar wieder deutlich hervor, besitzt aber dasselbe Vorzeichen wie in Luft oder Wasser. Ganz ähnlich verhalten sich sehr feine in Glas eingeritzte Gitter und ferner die eigentümlichen Gebilde, die man als Tonerdefasern bezeichnet [4]). Es sind dies Fasern, die aus metallischem Aluminium nach Benetzen mit Sublimatlösung hervor-

---

1) H. Ambronn, Kolloidzeitschr. **18**, 90 und 273, 1916; **20**, 173, 1917.

2) Ders., Zeitschr. f. Kristallogr. **52**, 48, 1913.

3) M. Schultze, Verhandl. d. naturhist. Ver. d. preuß. Rheinlande und Westfalens, **20**, 39, 1863.

4) H. Wislicenus, Kolloidzeitschr. **6**, 17, 87, 1910.

wachsen. Auch hier kann man die Doppelbrechung zum Verschwinden bringen, wenn man Flüssigkeiten von bestimmtem Brechungsexponenten eindringen läßt[1]). Kennzeichnend auch für alle diese Fälle ist, daß das Vorzeichen der Doppelbrechung von dem Vorzeichen der Differenz $n_1 - n_2$ unabhängig bleibt, wenn n_1 wieder den Brechungsexponenten der festen Substanz und n_2 den der Flüssigkeit bedeutet.

Man bezeichnete früher diese und ähnliche Fälle von optischer Anisotropie als Gitter- oder Lamellarpolarisation; eine genauere quantitative Untersuchung der Erscheinung war aber nicht ausgeführt worden. Erst in neuerer Zeit ist von O. Wiener[2]) eine allgemeinere Theorie des Verhaltens von Mischkörpern aufgestellt worden, die zeigt, daß in einem System aus zwei an sich völlig isotropen Komponenten Doppelbrechung auftreten kann, die von der Differenz $n_1 - n_2$ und von der Form und Anordnung der das System aufbauenden Teilchen abhängt. Von den beiden Grenzfällen, in denen die Doppelbrechung ein Maximum erreicht, ist für unsere Betrachtungen nur der von Wiener als Stäbchendoppelbrechung bezeichnete zu berücksichtigen. Haben die Teilchen, wenigstens der einen Komponente, Stäbchenform, d. h. sind sie räumlich in der Weise anisotrop, daß sie in zwei zu einander senkrechten Richtungen verschiedene Ausdehnung besitzen, und sind ihre Dimensionen klein gegen die Lichtwellenlänge, so gelten für völlige Parallelstellung aller Teilchen folgende Gleichungen:

$$n_0^2 - n_a^2 = -\frac{\delta_1 \delta_2 (n_1^2 - n_2^2)^2}{\delta_1 n_2^2 + \delta_2 n_1^2 + n_2^2}; \quad n_a^2 = \delta_1 n_1^2 + \delta_2 n_2^2; \quad \delta_1 + \delta_2 = 1$$

worin n_0 und n_a die Brechungsexponenten des ordentlichen und außerordentlichen Strahls im Mischkörper, n_1 und n_2 die Brechungsexponenten der beiden Komponenten und δ_1, δ_2 ihr Volumenverhältnis bezeichnen. Die Gleichungen gelten in dieser Form streng nur für farblose Bestandteile. Wie man sieht muß die Doppelbrechung stets positiv sein; denn $(n_1^2 - n_2^2)^2$ ist immer > 0 gleichgültig ob $n_1 - n_2 <$ oder > 0 wird, folglich muß $n_a > n_0$ sein. Wird $n_1 = n_2$, so muß Isotropie eintreten. Da dies aber im allgemeinen nicht für alle Farben gleichzeitig geschieht, besonders wenn die eine Komponente fest und die andere flüssig ist, so müssen ganz ähnliche Verhältnisse in der Dispersion der Doppel-

---

1) H. Ambronn, Kolloidzeitschr. 6, Heft 4, 1910.

2) O. Wiener, Ber. d. Kgl. Sächs. Ges. d. Wiss. math.-phys. Kl. 62, 255, 1910; 61, 113, 1909; Abhandl. ders. Ges. 32, Nr. 6, 1912.

brechung eintreten, wie wir sie für die schon erwähnten Fälle kennen. Es wird z. B. bei $(n_1)_D = (n_2)_D$ für die D-Linie zwar Isotropie bestehen, aber für die B- und die E-Linie muß der Mischkörper noch Stäbchendoppelbrechung zeigen, allerdings von demselben Vorzeichen. Während in diesem Gebiet die Reihenfolge der Chromozyklitfarben auftritt, müssen sich in dessen Nähe auch die anderen Farbenfolgen, wie sie oben schon charakterisiert wurden, bemerkbar machen. Also auch ohne ein Zusammenwirken zweier Anisotropieen von entgegengesetztem Vorzeichen finden sich bei der reinen stets positiven Stäbchendoppelbrechung schon jene starken Abweichungen von der normalen Reihenfolge der Interferenzfarben. Ich habe in einer früheren Arbeit hierfür ein im Einzelnen durchgerechnetes Zahlenbeispiel gegeben, auf das ich hier nur verweisen kann[1]).

3. Zusammenwirken von Stäbchendoppelbrechung und Eigendoppelbrechung.

Vergleicht man nun den Verlauf der Stärke der akzidentellen Doppelbrechung im Zelloidin, wie er im Absatz 1 dargestellt wurde, mit dem Gang der reinen Stäbchendoppelbrechung, so fällt vor allem die große Ähnlichkeit auf, die in der Abhängigkeit von der Differenz $n_1 - n_2$ besteht. Aber es ist doch auch ein wesentlicher Unterschied vorhanden. Beim Zelloidin wird für ein bestimmtes Gebiet der Werte von n_2 dieDoppelbrechung negativ, und sie verschwindet nicht für $n_1 = n_2$, sondern bei zwei Werten für n_2, von denen der eine kleiner und der andere größer als n_1 ist. Danach kann bei dem Verhalten des Zelloidins die positive Stäbchendoppelbrechung nicht allein wirksam sein, es muß sich ihr vielmehr noch eine andere Doppelbrechung von negativem Vorzeichen überlagern, die auf eine Eigenschaft der Zelloidinteilchen selbst zurückzuführen ist. Wir hätten also neben der Stäbchendoppelbrechung noch eine zweite Komponente zu berücksichtigen, nämlich die Eigendoppelbrechung, wie ich sie genannt habe. Aus dem Zusammenwirken dieser beiden Anisotropieen von entgegengesetztem Vorzeichen läßt sich nun der Verlauf der Gesamtdoppelbrechung, deren Stärke oben mit α_0 bezeichnet wurde, ganz ungezwungen erklären.

Handelt es sich dabei, wie angenommen werden darf, um eine einfache Überlagerung beider Anisotropieen, so kann man auch aus den Werten für α_0 die Stärke der Eigendoppelbrechung, die

1) Kolloidzeitschr. 18, 91, 1916.

weiterhin stets mit α_1 bezeichnet werden soll, bestimmen und somit zugleich den Gang der Stäbchendoppelbrechung rein darstellen; für die Stärke dieser stets positiven Komponente soll die Bezeichnung α_2 gewählt werden.

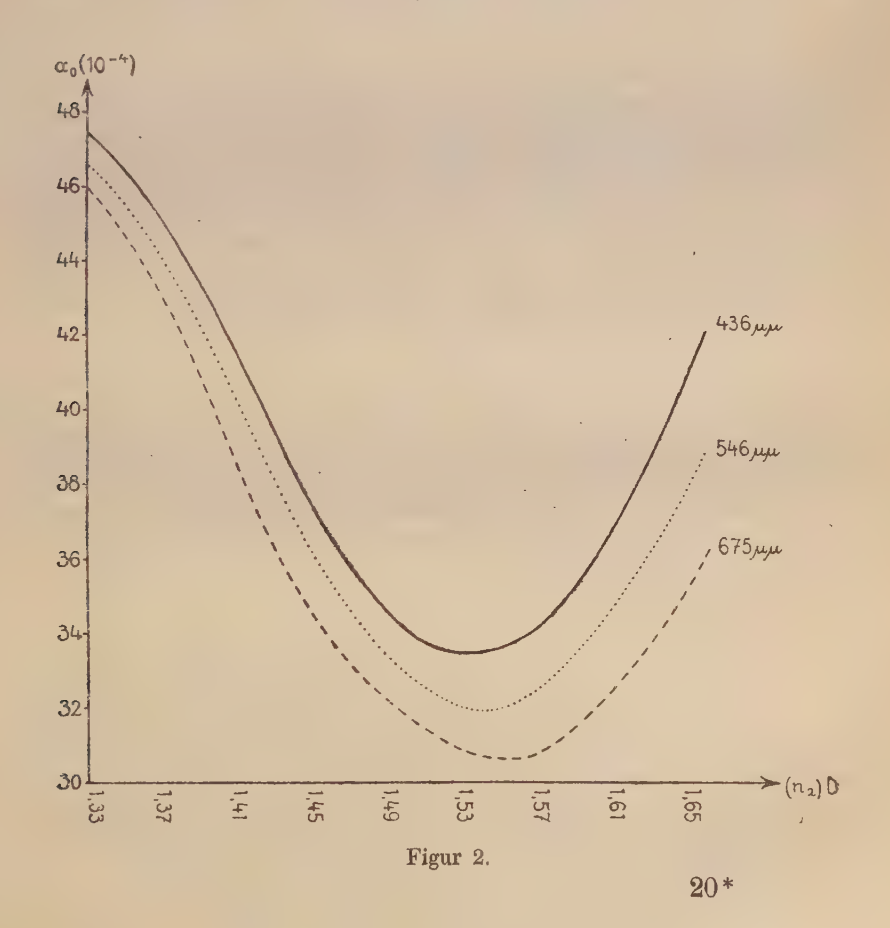
Indem man diejenigen Flüssigkeiten sucht, bei denen α_0 für die einzelnen Wellenlängen ein Minimum wird, weil durch Gleichheit von n_1 und n_2 jeweils α_2 verschwindet, und die Eigendoppelbrechung also rein hervortreten muß, so läßt sich unter Berücksichtigung des Anteils der Trockensubstanz der Streifen δ_1 und der Gesamtdicke der Wert für α_1 leicht berechnen. Kennt man aber α_0 und α_1, so kann man auch den Gang der Werte für α_2 rein darstellen. Führt man diese Berechnungen aus — wegen der Einzelheiten muß ich auch hierfür auf meine früheren Veröffentlichungen verweisen[1]) —, so zeigt sich in der Tat, daß der hier in Betracht kommenden Stäbchendoppelbrechung ebenfalls die theoretisch zu erwartenden großen Abweichungen in der Dispersion eigen sind. Gerade dieser Umstand aber scheint mir ein weiterer Beweis dafür zu sein, daß wir es in der einen Komponente der akzidentellen Doppelbrechung des Zelloidins mit Stäbchendoppelbrechung zu tun haben, die mit der Differenz $n_1 - n_2$ variabel ist, während die hinzukommende negative Eigendoppelbrechung konstant bleibt. Bei den in der Einleitung erwähnten Fällen der verschiedenen Varietäten von Apophylliten und der Mischkristalle von Blei- und Strontiumdithionat tritt im Zusammenwirken beider Komponenten nur einmal eine Umkehr des Vorzeichens und damit auch nur ein Nullwert für die Gesamtdoppelbrechung auf. Bei den deformierten Zelloidinstreifen findet dagegen zweimal eine Umkehr und damit auch zweimal ein Verschwinden der Gesamtdoppelbrechung statt. Während bei jenen echten Kristallen also jede der anomalen Reihenfolgen der Interferenzfarben nur einmal auftreten kann, wiederholen sich diese im Zelloidin mehrere Male und zeigen infolgedessen auch noch eine größere Mannigfaltigkeit.

4. Akzidentelle Doppelbrechung in der Zellulose.

Diese anomalen Reihenfolgen der Interferenzfarben müssen aber ganz wegfallen, wenn in einem solchen System auch die Eigendoppelbrechung positiv wird. Beim Zusammenwirken von Stäbchendoppelbrechung mit positiver Eigendoppelbrechung müssen die Werte für α_0 stets größer als 0 sein. Ein System mit dieser Eigenschaft kann man nun leicht herstellen, indem man die Zello-

1) Kolloidzeitschr. **18**, 278 und 279, 1916; **20**, 176 und 177, 1917.

idinstreifen durch geeignete Mittel wieder in Zellulose umwandelt. Wie ich früher gezeigt habe, kann dies geschehen, und zwar ohne wesentliche Änderung der äußeren Form der Streifen, wenn man sie längere Zeit mit Ammoniumsulfid behandelt und dann genügend mit reinem Wasser auswäscht. Die Stärke der Gesamtdoppelbrechung wird dadurch ganz beträchtlich erhöht. Aber auch in den Zellulosestreifen wirkt die Stäbchendoppelbrechung mit. Das geht daraus hervor, daß beim Verändern des Brechungsexponenten der Imbibitionsflüssigkeit sich ein ganz ähnlicher Gang der Werte für α_0 ergibt, nur daß diese Werte niemals bis auf 0 herabgehen oder gar kleiner als 0 werden. Steigert man wieder wie früher den Brechungsexponenten n_2 von 1,33 bis auf etwa 1,66, so zeigt die Beobachtung den Verlauf der Werte α_0, wie er für verschiedene Wellenlängen in beistehender graphischen Darstellung (Fig. 2) wiedergegeben worden ist. Auch bei der



Figur 2.

akzidentellen Doppelbrechung der Zellulose sinkt zunächst der Wert für α_0 infolge der Steigerung von n_2 beträchtlich, erreicht bei $n_2 = 1{,}54$ für mittlere Wellenlänge ein Minimum und steigt dann wieder an. Auch hieraus darf wieder der Schluß gezogen werden, daß sich der Stäbchendoppelbrechung eine Eigendoppelbrechung überlagert, die aber jetzt positiv ist. Stellt man nun wieder den Verlauf der Werte von α_2 in der vorhin schon angegebenen Weise rein dar, so ergibt sich auch hier die charakteristische Abweichung von dem normalen Gang der Dispersion. Die dadurch hervorgerufenen anomalen Reihenfolgen der Interferenzfarben können aber in den Zellulosestreifen nicht direkt beobachtet werden, sie treten nur im Verlauf der Werte für α_2 hervor und kommen auch in den Zahlen für α_0 zum Ausdruck, indem der Wert für den Quotienten $\frac{(\alpha_0)_{436\mu\mu}}{(\alpha_0)_{675\mu\mu}}$ von 1,08 bei $n_2 = 1{,}33$ allmählich bis auf 1,17 bei $n_2 = 1{,}66$ ansteigt. Dabei scheint mir die Tatsache wichtig, daß die Stäbchendoppelbrechung beim Zelloidin und bei der Zellulose, abgesehen von einigen kleineren Abweichungen, auf die hier nicht weiter eingegangen werden soll, auch quantitativ annähernd denselben Verlauf nimmt. Dies beweist, daß im Aufbau des ganzen Systems durch die Denitrierung des Zelloidins zur Zellulose keine wesentliche Änderung eingetreten ist. Die Stäbchenform und auch die Anordnung der Teilchen ist also dieselbe geblieben. Es war dies auch zu erwarten, wie aus den späteren Angaben über das Verhalten der natürlichen Zellhäute bei der Nitrierung und Denitrierung hervorgeht.

Wir haben es somit in der akzidentellen Doppelbrechung des Zelloidins wie der daraus gewonnenen Zellulose mit dem Zusammenwirken zweier Anisotropieen zu tun, und der Unterschied besteht nur darin, daß die mit dem Werte von n_2 veränderliche aber stets positive Stäbchendoppelbrechung beim Zelloidin von einer negativen und bei der Zellulose von einer positiven Eigendoppelbrechung überlagert wird. Die beiden Körper unterscheiden sich außerdem noch dadurch, daß der absolute Betrag von α_1 bei der Zellulose um mehr als das 10fache höher ist als beim Zelloidin, während die Werte für α_2, wie schon erwähnt, nur wenig von einander abweichen. So kommt es, daß die Gangunterschiede bei gleicher Dicke der Streifen im wasserimbibierten Zustande für dieselbe Wellenlänge bei der Zellulose etwa viermal größer sind als beim Zelloidin.

5. Optisches Verhalten der eingetrockneten Zelloidin- und Zellulosestreifen.

Noch eine Reihe anderer Beobachtungen spricht für die Richtigkeit der gemachten Annahme eines Zusammenwirkens von Stäbchen- und Eigendoppelbrechung. Läßt man nämlich die imbibierten Streifen in der Weise eintrocknen, daß dabei eine Verkürzung in der Längsrichtung vermieden wird, so erreichen die Unterschiede zwischen Zelloidin und Zellulose, wenigstens soweit die Werte für α_1 und α_2 in Betracht kommen, einen noch bedeutend höheren Betrag. Während des Eintrocknens wird die Dicke der Streifen sowohl beim Zelloidin wie bei der Zellulose um nahezu die Hälfte verringert. Der Gangunterschied beim Zelloidin vermindert sich dabei ebenfalls um die Hälfte, α_0 behält also ungefähr denselben Wert, wenigstens für eine mittlere Wellenlänge, und auch das positive Vorzeichen. Bei der Zellulose steigert sich dagegen der Gangunterschied trotz der erheblich geringeren Dicke um fast das Doppelte; der Wert für α_0 wird also dabei fast viermal größer als im imbibierten Zustande. Während die mit Wasser imbibierten Zelloidinstreifen die normale Reihenfolge der Interferenzfarben zeigen, treten in den eingetrockneten die Leukozyklit- oder gar die Andreasberger Farben auf; es ist also durch das Trocknen eine wesentliche Veränderung in der Dispersion der Gesamtdoppelbrechung bewirkt worden. In den Zellulosestreifen unterbleibt diese Änderung; sowohl im imbibierten wie im trocknen Zustand beobachtet man die gewöhnlichen Newtonschen Farben.

Die Erklärung für diesen zunächst etwas überraschenden Unterschied zwischen Zelloidin und Zellulose liegt ziemlich nahe, wenn man an der Annahme des Zusammenwirkens von Stäbchendoppelbrechung und Eigendoppelbrechung festhält. Die Verringerung des Querdurchmessers beim Eintrocknen wird jedenfalls eine Umlagerung der stäbchenförmigen Teilchen in dem Sinne herbeiführen, daß eine weitere Annäherung an die Parallelstellung ihrer Längsachsen eintritt. Dadurch muß nicht nur die Stäbchendoppelbrechung sondern auch die Eigendoppelbrechung erhöht werden. Die räumlichen Veränderungen kommen also beim Zelloidin wie bei der Zellulose in annähernd gleichem Maße in Betracht. Die Eigendoppelbrechnng des Zelloidins ist aber negativ und die der Zellulose positiv. Infolgedessen müssen die Stärken der Gesamtdoppelbrechung in den trocknen Streifen beider Körper einen wesentlichen Unterschied zeigen, was die Beobachtung auch bestätigt.

Leider läßt sich der Anteil der Eigen- und der Stäbchendoppelbrechung in den eingetrockneten Streifen nicht durch direkte

Messung feststellen, da es nicht möglich ist, diese mit verschieden brechenden Medien zu durchtränken und so die n_2-Werte zu suchen, für die die Gesamtdoppelbrechung jeweils ein Minimum erreicht. Wohl aber kann man auf einem Umwege dazu gelangen, wenigstens angenäherte Werte für α_1 und α_2 auch in diesen Fällen zu erhalten. Auf Grund der schon angedeuteten Überlegungen über die Art der Umlagerung der Teilchen, nach denen im Zelloidin und in der Zellulose durch das Trocknen ungefähr dieselben räumlichen Veränderungen eintreten, darf man folgende Annahmen machen: Es werde die Stärke der Eigendoppelbrechung um das p-fache und die der Stäbchendoppelbrechung um das q-fache erhöht. Bezeichnet man mit α_1' und α_1'' die Stärken der Eigendoppelbrechung der mit Wasser imbibierten Zelloidin- und Zellulosestreifen, so wären danach die entsprechenden Werte für die trocknen Streifen $p\alpha_1'$ und $p\alpha_1''$; und in ähnlicher Weise könnte man für die Stärken der Stäbchendoppelbrechung in den trocknen Streifen die Werte $q\alpha_2'$ und $q\alpha_2''$ bilden. Die Zahlen p und q kann man aus einer Reihe von Beobachtungen berechnen; denn es ergeben sich für sie folgende Gleichungen:

$$\alpha_0 = \delta_1 p\alpha_1' + q\alpha_2' \text{ und } \alpha_0^* = \delta_1^* p\alpha_1'' + q\alpha_2''$$

in denen α_0 die Stärke der Gesamtdoppelbrechung der trocknen Zelloidin- und α_0^* diejenige der trocknen Zellulosestreifen bezeichnen, während δ_1 und δ_1^* die mit Rücksicht auf die Querschnittverringerung zu bestimmenden Anteile der festen Komponenten angeben. Man erhält aus dieser Berechnung die Mittelwerte $p = 2{,}3$ und $q = 1{,}6$. Die Erhöhung der Eigendoppelbrechung ist also beträchtlich größer, als die der Stäbchendoppelbrechung. Für die Verstärkung der ersteren kommt wohl nur die Annäherung an die Parallelstellung der Teilchen in Betracht, während für die Veränderung der Stäbchendoppelbrechung mehrere Ursachen wirksam werden können. Außer der Annäherung an die Parallelstellung, die zugleich eine Annäherung an den Grenzfall der reinen Stäbchendoppelbrechung bedeutet, ist noch die Veränderung des Anteils δ_1 an dem Aufbau des Systems wichtig. Im imbibierten Zustande beträgt dieser Wert etwa 0,25, während er im trocknen Zustande in der Nähe von 0,5 liegt. Nun geht aus der Wienerschen Gleichung hervor, daß die Stärke der Stäbchendoppelbrechung bei Annäherung an diesen Wert ein Maximum erreicht. Schließlich könnte durch das Eintrocknen auch noch die Differenz $n_1 - n_2$ geändert werden; auch hiermit würde, wenn der Wert dieser Differenz wächst, eine Erhöhung von α_2 verbunden sein. Die bisherigen

Untersuchungen lassen nicht mit Sicherheit erkennen, ob im lufttrocknen Zustande noch geringe Flüssigkeitsmengen zwischen den Stäbchen zurückbleiben, oder ob die Zwischenräume einen gasförmigen Körper enthalten. Im letzteren Falle würde hierin wohl die Hauptursache für die Erhöhung der Stäbchendoppelbrechung zu suchen sein.

Genauere Zahlenangaben über diese durch das Eintrocknen der Streifen verursachten Änderungen der Werte α_0, α_1, α_2 sollen an anderer Stelle veröffentlicht werden. Hier möchte ich nur kurz auf zwei Punkte aufmerksam machen. Die Stäbchendoppelbrechung wird in beiden Fällen ungefähr denselben Wert behalten, da sie im imbibierten Zustande ebenfalls für Zelloidin und Zellulose nahezu gleich ist. Dagegen wird die Eigendoppelbrechung ganz beträchtlich verstärkt, und dabei zeigt sich ein sehr bemerkenswerter Unterschied zwischen dem Zelloidin und der Zellulose. Die Untersuchungen, die ich mit einem Zelloidin ausführte, das vor dem Kriege bezogen worden war, ergaben in einem bestimmten Punkte ein anderes Resultat; wie die an einem aus dem Herbst 1917 stammenden Zelloidin. Der Unterschied tritt aber erst nach der Denitrierung der Streifen auf. Beide Zelloidinsorten ergaben sowohl für α_0 wie für α_1 und α_2 nahezu dieselben Werte. Die denitrierten Streifen verhielten sich dagegen insofern verschieden, als bei den alten Präparaten, für deren Herstellung reine Baumwolle verwendet worden war, die Werte α_0 und α_1 beträchtlich kleiner waren, als bei den Präparaten aus dem Jahre 1917. Nach freundlicher Auskunft der Fabrik (Chem. Fabrik auf Aktien, vorm. E. Schering, Berlin.) ist zu diesem Zelloidin keine Baumwolle sondern ein ganz anderer Faserstoff als Rohstoff verwendet worden. Während also die aus beiden Faserstoffen hergestellte Nitrozellulose sich fast ganz gleich verhält, zeigt die daraus gewonnene Zellulose beträchtliche Verschiedenheiten je nach dem Ausgangsmaterial. Es ist das nicht gerade überraschend, da man weiß, daß verschiedene Zellulosehäute auch verschieden starke Doppelbrechung besitzen. Diese ursprüngliche Verschiedenheit im optischen Verhalten scheint nun trotz den tiefeingreifenden Prozessen der Nitrierung und Denitrierung erhalten zu bleiben, wenigstens für die Zellulose, während die Stärke der Eigendoppelbrechung der Nitrozellulose im wesentlichen von der Höhe des Stickstoffgehaltes abhängt und die Natur des Ausgangsmaterials hierauf so gut wie keinen Einfluß ausübt. Ich werde später auf diesen Punkt nochmals zurückkommen.

6. Optisches Verhalten der Nitrozellulosefasern.

Im Vergleich zu den im Vorstehenden dargelegten Ergebnissen ist es nun von Interesse, das Verhalten der nicht zu Zelloidin verarbeiteten Nitrozellulosefasern zu betrachten. Es kann hierüber nur ein ganz kurzer Überblick an dieser Stelle gegeben werden. Schon vor mehr als 50 Jahren hatte C. Nägeli[1]) gefunden, daß die Doppelbrechung der Schießbaumwolle bei schwacher Nitrierung noch das positive Vorzeichen, bei höherem Stickstoffgehalt jedoch das negative Vorzeichen besitzt, während die äußere Form der Fasern gar nicht verändert erscheint. Diese Umkehr des Charakters der Doppelbrechung ließ schon erwarten, daß auch eine Nitrierungsstufe bestehe, bei der die Fasern wenigstens für eine Wellenlänge optisch isotrop werden. Von Nägeli selbst ist diese auffallende Erscheinung nicht weiter verfolgt worden, und seine Beobachtungen scheinen überhaupt ganz in Vergessenheit geraten zu sein. Erst neuere systematisch durchgeführte Untersuchungen von Hans Ambronn[2]) haben mit Sicherheit ergeben, daß man von der positiven Doppelbrechung der reinen Zellulosefasern bis zu dem Verhalten der höchsten Nitrierungsstufen, die für alle Farben negative Doppelbrechung zeigen, eine Reihe von Übergängen beobachten kann, und daß diese sowohl beim Vorgange der Nitrierung wie der Denitrierung auftreten. Zugleich weist die Dispersion der Doppelbrechung dieselben Abweichungen in der Reihenfolge der Interferenzfarben auf, wie sie bei den Apophylliten und den Mischkristallen aus Blei- und Strontiumdithionat bestehen. Um dieses Verhalten übersichtlich zu demonstrieren braucht man nur in einem Präparate verschiedene Fasern derselben Herkunft so nebeneinander zu legen, daß neben der reinen Zellulosefaser mehrere Nitrierungsstufen von steigendem Stickstoffgehalt zugleich beobachtet werden können. Man übersieht dann in einem Bilde, wie die reine Faser stark positiv doppelbrechend ist, während in den niederen Nitrierungsstufen die Stärke der, zunächst noch positiven, Doppelbrechung allmählich geringer wird, bei einem bestimmten Stickstoffgehalt, für eine mittlere Wellenlänge, ganz verschwindet und von da ab wieder ansteigt, nun aber das negative Vorzeichen erhält. Betrachtet man die verschiedenen Fasern im gewöhnlichen Lichte, so lassen sie sich überhaupt kaum unterscheiden; die äußere Form ist also trotz den starken Verschieden-

---

1) C. Nägeli, Ber. d. Bayr. Akad. I, 1862, 307.

2) Hans Ambronn, Über die Änderung d. opt. Verhaltens d. Zellulose bei d. Nitrierung. (Diss. Jena, 1914.)

heiten im chemischen und physikalischen Verhalten so gut wie unverändert geblieben. Diese Abhängigkeit der optischen Eigenschaften von der Höhe des Stickstoffgehalts erinnert an das eigentümliche Verhalten der von F. Pockels untersuchten Flintgläser, bei denen der Charakter und die Stärke der akzidentellen Doppelbrechung von der Höhe des Bleigehalts abhängt, wie in der Einleitung bereits erwähnt wurde.

Das zunächst etwas überraschende Verhalten der Fasern läßt sich jedoch auf Grund der von mir schon früher gemachten Annahme verstehen, daß die einzelnen stäbchenförmigen Teilchen der Zellulose, die Micelle nach Nägeli[1]), durch die Nitrierung in ihrer Form nicht wesentlich geändert werden, sondern nur einen chemischen Umbau erfahren, und daß ferner bei der Denitrierung ebenfalls unter Beibehaltung des rein morphologischen Aufbaus diese chemische Umwandlung wieder vollständig rückgängig gemacht wird. Bleibt aber in beiden Fällen die räumlich anisotrope Gestalt der Micelle erhalten, so kann natürlich auch die ganze Faser ihre Form behalten, und es braucht keine Veränderung in der gegenseitigen Anordnung der stäbchenförmigen Micelle einzutreten; vorausgesetzt daß der chemische Umbau weder bei der Nitrierung noch bei der Denitrierung unter stürmischen Vorgängen verläuft, die eine Desorientierung der Micelle herbeiführen. Wird aber nun bei der Denitrierung das Zellulosemicell mit allen seinen früheren Eigenschaften wieder hergestellt, so muß es auf das Gesamtverhalten der Membran auch Einfluß haben, von welcher Fasersorte man ausgegangen war; denn die Stärke der Doppelbrechung ist durchaus nicht bei allen Fasersorten dieselbe. Auch die oben erwähnten Verschiedenheiten zwischen Zelloidinen oder vielmehr den daraus gewonnenen Zellulosen aus verschiedenem Rohmaterial dürften wohl durch diesen Umstand ihre befriedigende Aufklärung finden.

7. Nägeli's Micellarhypothese.

Wie man sich nun im Einzelnen den Vorgang der Nitrierung zu denken hat, ob dabei die Zellulosemicelle bei schwächerer Nitrierung nur zum Teil umgebaut werden und erst bei stärkerer Einwirkung eine völlige Umwandlung erfahren, oder ob in den verschiedenen Nitrierungsstufen chemisch ganz von einander abweichende Nitrozellulosemicelle entstehen, das ist eine mehr chemische Frage, die hier nicht weiter berührt zu werden braucht.

---

1) C. Nägeli, Theorie d. Gärung, 121, 1879.

Es genügt, darauf hinzuweisen, daß im optischen Verhalten jedenfalls alle Übergänge von der reinen Zellulose bis zur Nitrozellulose mit sehr hohem Stickstoffgehält zu beobachten sind.

Nach der Nägelischen Micellarhypothese[1]) hat man sich die Micelle als Komplexe von chemischen Molekülen zu denken, die aus Hunderten ja Tausenden von Einzelmolekülen aufgebaut sind. Es darf deshalb auf Grund dieser Hypothese ohne Bedenken angenommen werden, daß bei chemischen Eingriffen ebensowohl eine rasche Umwandlung der ganzen Micelle wie ein allmählich erfolgender Ab- und Umbau stattfinden kann. Der letztere Vorgang dürfte sogar in manchen Fällen größere Wahrscheinlichkeit für sich haben, denn gerade in der neueren physiologischen Chemie kennt man zahlreiche Beispiele für einen solchen Schritt für Schritt verlaufenden Ab- und Umbau der Teilchen ähnlicher Körper, für die ihrem kolloidalen Charakter nach ebenfalls eine micellare Struktur im Sinne Nägelis anzunehmen ist.

Es mag noch kurz darauf hingewiesen werden, daß Nägeli zu der Aufstellung des Begriffs „Micell" durch seine Untersuchungen über den Bau der Stärkekörner gekommen ist und zwar hauptsächlich durch die Versuche über die Quellungserscheinungen dieser Gebilde. Die regelmäßigen Ungleichwertigkeiten der Richtungen, die sich dabei zeigten, veranlaßten ihn, die Teilchen als polyedrisch geformt mit verschiedener Ausdehnung nach verschiedenen Richtungen anzunehmen; er verglich sie deshalb direkt mit kristallinischen Körpern. In den ersten Veröffentlichungen bezeichnete er diese Teilchen noch als Moleküle; später als dieser Begriff in der Chemie einen ganz anderen Inhalt erhielt, nannte er sie Micelle (micellum, Dimin. von mica.), und die von ihm gemachte Annahme einer kristallinischen Beschaffenheit schien ihm durch das Verhalten der Stärkekörner und Zellhäute im polarisierten Lichte eine willkommene Bestätigung zu finden. Denn denkt man sich eine sehr große Anzahl von Micellen mit ihren gleichsinnigen Achsen parallel oder doch annähernd parallel gelagert, so müssen sich ihre Einzelwirkungen summieren und das ganze System muß dann eine starke optische Anisotropie zeigen, wenn auch die regelmäßige Ausbildung geometrisch gut charakterisierter Flächen, wie bei den echten Kristallen, dabei keineswegs zustande zu kommen braucht. Die Nägelische Auffassung kommt also ungefähr auf dasselbe hinaus wie die Langevinsche Orientierungshypothese, die zuerst für die Erklärung der elektro- und

1) C. Nägeli, Die Stärkekörner, 332, 1858.

magnetooptischen Doppelbrechung mancher kolloider Lösungen aufgestellt wurde, und die seitdem noch für eine Reihe anderer ähnlicher Erscheinungen im Gebiete der Kolloidforschung recht befriedigende Aufklärung ergeben hat.

Wendet man nun die Nägelische Micellarhypothese auf die Erscheinungen im Zelloidin und in der daraus gewonnenen Zellulose an, so erhält man ebenfalls eine ganz plausible Erklärung für den in mancher Beziehung überraschenden Verlauf der akzidentellen Doppelbrechung in diesen Körpern. Sind Zelloidin und Zellulose aus Micellen aufgebaut, die räumlich und optisch an sich schon Anisotropie zeigen, so wird durch die einseitige starke Deformation eine Orientierung der stäbchenförmigen Micelle bewirkt. Bestehen die festen Teilchen des Zelloidins aus einer Nitrozellulose, die negative Doppelbrechung besitzt, so wird die Summe der Einzelwirkungen in den gedehnten Streifen allein ebenfalls eine negative Gesamtdoppelbrechung aufweisen, die aber nur dann rein hervortreten kann, wenn die flüssige Komponente des Systems denselben Brechungsexponenten besitzt wie die festen Teilchen selbst. Ist dagegen das Brechungsvermögen beider Komponenten verschieden, so muß dieser negativen Eigendoppelbrechung eine Stäbchendoppelbrechung überlagert werden, deren Vorzeichen stets positiv ist, gleichgültig ob die Differenz $n_1 - n_2$ positiv oder negativ wird. Je höher der absolute Wert von $n_1 - n_2$ ist, desto größer muß auch die Stärke dieser Stäbchendoppelbrechung sein. Somit muß bei Steigerung des Wertes von n_2 von 1,33 bis über 1,7 hinaus der Verlauf der Gesamtdoppelbrechung des ganzen Systems von positiv durch 0 nach negativ und von da ab wieder durch 0 nach positiv sich in der Weise gestalten, wie es die Beobachtungen mit voller Sicherheit ergeben haben (vergl. Fig. 1). Und besitzen anderseits die räumlich nahezu gleichen Zellulosemicelle positive Eigendoppelbrechung, so muß sich die Summe ihrer Einzelwirkungen ebenfalls zu der Stäbchendoppelbrechung addieren und somit eine stets positive Gesamtdoppelbrechung ergeben, die bei Variierung von n_2 niemals auf 0 herabgehen, sondern nur ein Minimum erreichen kann, das bei $n_1 - n_2 = 0$ eintritt. Der Verlauf muß also in der Weise sich gestalten, daß wiederum unabhängig von dem Vorzeichen der Differenz $n_1 - n_2$ die Stärke der Gesamtdoppelbrechung mit dem absoluten Wert dieser Differenz auf beiden Seiten des Minimums ansteigt, was ebenfalls mit den Beobachtungen vollkommen übereinstimmt (vergl. Fig. 2). Auch das in der Einleitung erwähnte Verhalten des Zelluloids steht mit diesen Ergebnissen keineswegs im Widerspruch, denn die zur

Herstellung dieses Körpers verwendete Nitrozellulose stellt eine niedrigere Nitrierungsstufe dar, deren Micelle noch positive Eigendoppelbrechung besitzen. Hierdurch findet die nach dem Herauslösen des Kampfers noch fortbestehende positive Gesamtdoppelbrechung eine befriedigende Erklärung.

8. Dichroismus.

Die Ergebnisse über den Verlauf der akzidentellen Doppelbrechung im Zelloidin und in der Zellulose können als eine Bestätigung der Nägelischen Micellarhypothese angesehen werden. Zugleich stellen sie einen bemerkenswerten Fall von Stäbchendoppelbrechung im Zusammenwirken mit Eigendoppelbrechung dar. Beide Komponenten der untersuchten Systeme sind farblos, und die oben angeführte Form der Wienerschen Gleichungen ist deshalb anwendbar. Etwas anders gestaltet sich die Sache jedenfalls, wenn eine Komponente gefärbt ist. Theoretisch läßt sich zur Zeit nicht viel darüber sagen, was dann zu erwarten ist; es bleibt also nur übrig, durch Versuche festzustellen, wie sich das optische Verhalten unter diesen Umständen gestaltet.

Wie ich schon in mehreren früheren Untersuchungen nachweisen konnte[1]), ergibt sich bei Einführung von Farbstoffen in natürliche Fasern in vielen Fällen ein starker Dichroismus. Ganz ähnliche Verschiedenheiten in der Absorption treten nun auch an den gedehnten Zelloidin- und Zellulosestreifen auf. Färbt man z. B. Nesselfasern mit Kongorot, einem in der Mikrotechnik viel verwendeten Farbstoff, so erhält man einen starken Dichroismus in der Weise, daß die Fasern bei Parallelstellung ihrer Längsachsen mit der Polarisationsebene des Nicols fast ungefärbt, dagegen bei Kreuzung beider Richtungen dunkelrot erscheinen. Den gleichen Dichroismus zeigen auch gedehnte Zelloidinstreifen, wenn sie mit demselben Farbstoff gefärbt werden. Da die Färbung in konzentriertem Glyzerin vom Brechungsexponenten 1,46 und auch in geeigneten höher brechenden Flüssigkeiten unverändert bleibt, so kann man prüfen, wie sich der Dichroismus verhält, wenn die Streifen mit solchen Medien durchtränkt werden. Ungefärbte Streifen sind in diesen Flüssigkeiten für eine mittlere Wellenlänge optisch isotrop oder sie werden bei höherem Brechungsvermögen sogar für alle Farben negativ doppelbrechend. Trotzdem

---

1) H. Ambronn, Wiedem. Ann. **34**, 340, 1888, Arch. f. d. ges. Physiol. **44**, 301, 1889; Ber. d. Deutsch. Bot. Ges. **6**, 226, 1888 u. **7**, 103, 1889 Ber. d Kgl. Sächs. Ges. d. Wiss. math.-phys. Kl. **48**, 613, 1896.

bleibt in den gefärbten Streifen der Dichroismus in derselben Stärke erhalten wie bei der Durchtränkung mit Wasser. Es ist also die Lage der Achsen des Absorptionsellipsoids von der Stärke und dem Vorzeichen der Doppelbrechung in den Streifen vollkommen unabhängig. Maßgebend scheint nur die durch die Deformation bewirkte Orientierung der Micelle zu sein und zugleich die Eigenschaft der Farbstoffteilchen selbst. Sind diese schon an sich anisotrop und dichroitisch, so müssen sie bei annähernd gleichsinniger Einlagerung in die Streifen einen starken Gesamtdichroismus als Summe ihrer Einzelwirkungen hervorrufen, wie es die Beobachtung zeigt. Dieser Dichroismus kann aber auch dann nicht verschwinden, wenn die Streifen selbst isotrop oder negativ doppelbrechend werden, denn die einmal vorhandene Orientierung der Kongorotteilchen wird dadurch nicht verändert.. Untersucht man aber nun die mit konzentriertem Glyzerin imbibierten gefärbten Streifen, so erweisen sie sich nicht als optisch isotrop, sondern als stark doppelbrechend für rotes Licht. Diese Doppelbrechung kann also nur von den in dem System eingelagerten und orientierten Kongorotteilchen herrühren.

Durch Einwirkung einer schwachen Säure kann man das eingelagerte Kongorot in den entsprechenden blauen Farbstoff umwandeln und erhält auch dann wieder einen starken Dichroismus dunkelblau—blaßblau. Dadurch wird also ebenfalls ein Umbau der roten Teilchen in die blauen bewirkt, ohne daß eine Änderung in der Orientierung eintritt. Man kann unter dem Mikroskop gut verfolgen, wie bei langsamen Zufließen der Säure die Umfärbung allmählich längs der ganzen Faser fortschreitet; und es ist von Interesse, daß dabei in der schmalen Zone zwischen rot und blau, wo offenbar die Änderung des in fester Form eingelagerten Farbstoffs erfolgt, eine nicht dichroitische schmutzig gelbliche Färbung auftritt, die wie ein Ring über die ganze Faser hinwegzieht.

Von den zahlreichen anderen Färbungen, die starken Dichroismus ergeben, sollen hier nur noch wenige besonders charakteristische Beispiele angeführt werden. Behandelt man Zellulosemembranen mit Chlorzinkjodlösung, so färben sie sich dunkelviolett. Betrachtet man sie im polarisierten Licht, so erscheinen sie bei Parallelstellung ihrer Längsachsen mit der Polarisationsebene fast farblos, bei Kreuzung beider Richtungen dagegen völlig schwarz; sie verhalten sich also ähnlich wie Turmalin oder Herapathit. Auch die künstlich durch Denitrierung aus gedehntem Zelloidin gewonnenen Zellulosestreifen zeigen dasselbe Verhalten. Stellt

man zwischen fest aufeinandergepreßten Glasplatten sehr dünne Jodkriställchen her, die eine Dicke unter 0.1 μ besitzen, so zeigen diese fast ganz denselben starken Dichroismus, wie die mit Jod gefärbten Membranen. Auch daraus darf wohl geschlossen werden, daß hierbei der färbende Körper, mag dies nun Jod selbst oder eine Jodverbindung sein, in orientierter kristallinischer Form eingelagert wird.

Besonders wichtig scheinen mir auch manche Färbungen mit Metallen zu sein. Behandelt man in geeigneter Weise Fasern mit Goldchlorid oder Silbernitrat, so erhält man sehr intensive und stark dichroitische Färbungen. Mit Gold gefärbte Fasern erscheinen bei Parallelstellung von Faserachse und Polarisationsebene intensiv weinrot und bei Kreuzung tief blaugrün, bei mit Silber gefärbten Fasern sind die entsprechenden Farben schwefelgelb und tief indigoblau. Untersucht man diese Färbungen im monochromatischen Licht, so erkennt man sofort, daß die goldgefärbten Fasern im blaugrünen Licht bei Parallelstellung dunkel und in der Orthogonalstellung hell erscheinen, während im roten Lichte gerade das Umgekehrte eintritt und bei Beobachtung im gelben Lichte überhaupt kaum eine Verschiedenheit in der Absorption sich bemerkbar macht. Ganz entsprechend sind die Verhältnisse in den silbergefärbten Fasern, hier werden die gelben Strahlen parallel und die blauen senkrecht zur Faserachse polarisiert, und die Verschiedenheit der Absorption verschwindet fast ganz im grünen und roten Lichte. Berücksichtigt man nun, daß auch an Teilchen des kolloidalen Goldes und Silbers ganz ähnliche Verschiedenheiten in der Absorption nachgewiesen worden sind, so dürfte die Annahme nicht unberechtigt sein, daß die in den Fasern eingelagerten Gold- und Silberteilchen an sich schon eine Verschiedenheit der Absorption in verschiedenen Richtungen besitzen und durch die orientierte Einlagerung zwischen oder auf den Micellen den starken Dichroismus durch Summierung ihrer Einzelwirkungen herbeiführen. Für die Berechtigung dieser Auffassung scheinen mir auch Versuche zu sprechen, die ich in Gemeinschaft mit R. Zsigmondy[1]) ausgeführt habe und die eine Färbung von gedehnten Gelatinestreifen mit kolloiden Gold- und Silberlösungen zum Ziele hatten. Färbt man flüssige Gelatine mit derartigen Lösungen und läßt sie dann zu Platten eintrocknen, so kann man daraus Streifen schneiden und diese einer starken

---

1) H. Ambronn und R. Zsigmondy, Ber. d. Kgl. Sächs. Ges. d. Wiss. math.-phys. Kl. 51, 13, 1899.

Deformation unterwerfen. Sie zeigen dann ebenfalls einen Dichroismus, der dem der Fasern sehr ähnlich ist. Hierbei werden also auch die Gold- und Silberteilchen durch die starke Dehnung orientiert, und ihre Gesamtwirkung läßt darauf schließen, daß auch den einzelnen Teilchen an sich schon eine Verschiedenheit der Absorption zukommt.

Ähnlich wie man das Kongorot durch chemischen Eingriff innerhalb der Faser in einen blauen Farbstoff verändern kann, so kann man nun auch bei den Metallfärbungen verschiedene Umwandlungen hervorrufen. So läßt sich mit Sublimat das Silber in Amalgam überführen, die Faser zeigt dann einen Dichroismus braun—blaugrau; ferner kann man in der Faser zunächst das farblose Schwefelzink einlagern und dann mit Gold-, Silber- oder auch anderen Metallsalzen die entsprechenden Sulfide herstellen, die fast alle auch deutlichen Dichroismus hervorrufen[1]). Bei Schwefelgold läßt sich dann durch Erwärmung leicht eine reine Goldfärbung mit dem Dichroismus rot—blaugrün erzielen. In ähnlicher Weise kann man in der Faser zunächst eine sehr intensive und stark dichroitische Quecksilberfärbung durch abwechselnde Behandlung mit Sublimat und Zinnchlorür erreichen und dann mit Goldchlorid nachbehandeln. Sofort nach Einwirken des Goldsalzes erscheint die Faser in der reinen Goldfärbung mit dem charakteristischen Dichroismus rot—blaugrün, da durch das Quecksilber das Goldchlorid reduziert wird. Noch eine ganze Reihe anderer Reaktionen lassen sich in diesen intermicellaren Räumen hervorrufen, auf die hier nicht weiter eingegangen werden soll. Jedenfalls zeigen alle diese chemischen Umwandlungen das Gemeinsame, daß dabei ganz ähnlich wie bei den Vorgängen der Nitrierung und Denitrierung die vorher bereits vorhandene Orientierung der Teilchen nicht weiter geändert wird, also der Aufbau des ganzen Systems erhalten bleibt. Daß der Aufbau der Fasern oder gedehnten Streifen dabei nicht gestört wird, solange wenigstens nicht Zellulose- oder Zelloidinzerstörende Reagentien angewandt werden, darf wohl ohne weiteres angenommen werden. Somit muß aber in diesem Aufbau die für die orientierte Einlagerung der färbenden Teilchen bestimmende Ursache liegen. Hält man mit Nägeli an der Kristallnatur der diese Systeme aufbauenden Micelle fest, so läßt sich auch für das orientierte Einlagern gleichfalls kristalli-

---

1) H. Ambronn, Wiesner-Festschrift, 193, 1909. Vergl. auch: Emich und Donau, Über die Anwendung von Gespinstfasern in der mikrochemischen Analyse anorganischer Stoffe, Festschrift für Adolf Lieben, 651—663, 1906.

nischer Teilchen eine befriedigende Erklärung geben. Ähnlich wie auf frischgespaltenen Kristallflächen andere Körper orientiert auskristallisieren, so werden auch die von den Micellen ausgehenden richtenden Einflüsse die orientierte Ein- oder Auflagerung jener färbenden Teilchen, die an sich ebenfalls kristallinisch sind, herbeiführen und auch bei chemischen Umwandlungen die einmal bestehende Anordnung erhalten können. Auf Grund dieser Überlegungen glaube ich, daß man auch das ganze Gebiet dieses künstlichen Dichroismus und der dabei zu beobachtenden Erscheinungen als eine Bestätigung der Nägelischen Hypothese ansehen darf.

Jena, 1. März 1919.

Zusammenfassung der wichtigsten Ergebnisse:

1. Die akzidentelle Doppelbrechung des Zelloidins und der daraus gewonnenen Zellulose ist abhängig von dem Brechungsvermögen der Flüssigkeiten, mit denen diese Körper durchtränkt sind.
2. Bei derselben bleibenden Verlängerung durch Dehnung kann im Zelloidin die Doppelbrechung sowohl positiv, als negativ, als auch Null sein; in der Zellulose ist sie dagegen stets positiv und wird niemals Null.
3. Zur Erklärung dieser Versuchsergebnisse wird ein Zusammenwirken der Stäbchendoppelbrechung im Sinne O. Wieners mit der den stäbchenförmigen Teilchen zukommenden Eigendoppelbrechung angenommen; diese Eigendoppelbrechung hat beim Zelloidin negatives, bei der Zellulose positives Vorzeichen.
4. Die Erscheinungen der akzidentellen Doppelbrechung in beiden Körpern stellen nicht nur einen bemerkenswerten Fall von Stäbchendoppelbrechung dar, sondern können auch als eine Bestätigung der Nägelischen Theorie der kristallinischen Micelle angesehen werden.
5. Der starke Dichroismus, den sowohl natürliche Fasern, als auch die bleibend gedehnten Zelloidin- und Zellulosestreifen bei geeigneter Färbung zeigen, spricht ebenfalls für die Richtigkeit der Nägelischen Auffassung.

Untersuchungen aus dem Universitätslaboratorium zu Göttingen. XXX.

Von

O. Wallach.

Vorgelegt in der Sitzung vom 18. Juli 1919.

I. Über die Pulegonsulfosäure.

Schon vor längerer Zeit ist gefunden[1]), daß Pulegon (I) in Berührung mit wässriger schwefeliger Säure die Bestandteile der Säure aufnimmt, um eine in Wasser sehr leicht lösliche Sulfosäure zu liefern. Das ganz analog gebaute, nur statt eines 6-Rings einen 5-Ring enthaltende Kampferphoron (II) verhält sich ebenso

CH_3 I. $=O$ C CH_3 CH_3 CH_3 II. $=O$ CH_3 CH_3

Sehr beschleunigen kann man die Bildung der Pulegonsulfosäure, wenn man das Keton mit konzentrierter wässriger schwefeliger Säure auf der Schüttelmaschine durchschüttelt. Die so erhaltenen Lösungen sind aber relativ verdünnt und müssen durch

1) Ann. d. Chem. **331**, 332 (1904).

Absaugen von der schwefeligen Säure befreit werden[1]). Eine nur wenig Wasser enthaltende Pulegonsulfosäure kann man bereiten, wenn man eine Lösung von 1 Th. Pulegon in etwa 3 Th. absolutem Alkohol mit gasförmiger schwefeliger Säure sättigt und in lose verschlossenen Gefäßen einige Tage bei etwa 20° stehen läßt. Das freie Pulegon ist dann ganz verschwunden. Gießt man nun die Flüssigkeit in eine flache Schale und läßt den Alkohol bei gewöhnlicher Temperatur (ev. unter Druckverminderung) verdunsten, so hinterbleibt nach Entfernung des Lösungsmittels die Sulfosäure in Form eines fast farblosen Syrups[2]). Dadurch daß man diesen über konzentrierte Schwefelsäure stellt, kann man die Säure nicht wasserfrei erhalten, denn es tritt dabei Dunkelfärbung und Zersetzung unter Abspaltung von schwefeliger Säure und Pulegon ein.

Sehr charakteristisch ist für die Säure ihre Fähigkeit sehr gut krystallisierte Salze zu bilden, die sich im Gegensatz zu der wasserfreien Säure durch große Beständigkeit auszeichnen. Man erhält diese Salze am besten, wenn man eine nicht zu verdünnte wässrige Lösung der syrupösen Säure mit Metallkarbonaten neutralisiert, filtriert und bei gewöhnlicher Temperatur zur Krystallisation stellt[3]). Die Salze von Magnesium, Zink und Blei krystallisieren in schön ausgebildeten, in Wasser sehr leicht, in Alkohol etwas weniger löslichen Prismen. Ebenso das etwas schwerer lösliche Baryumsalz. Das Calciumsalz bildet Blättchen. Auch das Ammoniaksalz ist kristallinisch und recht beständig. Man erhält es durch Abdunsten der ammoniakalischen Lösung der Säure an der Luft. Entsprechend kann man Salze der organischen Amine darstellen, z. B. das auch kristallinische, beständige Piperidinsalz. Am wenigsten beständig sind die Alkalisalze, die man durch Umsetzung des Baryum- oder Calciumsalzes mit Alkalikarbonat gewinnen kann. Dabei ist aber ein Überschuß von Karbonat zu vermeiden, denn während alle Salze der Sulfosäure in trockenem Zustand und — abgesehen von den Alkalisalzen — auch in heißer wässriger Lösung sehr beständig sind, werden sie durch Ätzalkali (selbst Kalkwasser), sowie durch kohlensaures

---

1) Eine hydrolytische Spaltung des Pulegons in Methylcyclohexanon und Aceton bewirkt die schwefelige Säure nicht.

2) Die alkoholische Lösung enthält auch immer eine kleine Menge einer krystallinischen Substanz, die noch nicht untersucht ist. Vielleicht liegt Menthol vor (?).

3) Man kann die Salze auch gewinnen, indem man Metalle, wie Magnesium Zink, Calcium, in die wässrige Lösung einträgt. Es bietet das aber keine Vorteile.

Alkali schon bei schwacher Temperatursteigerung mehr oder weniger vollständig unter Abscheidung von gewöhnlichem Pulegon (nicht Isopulegon) zersetzt. Dies Verhalten kann man bei der Analyse nutzbar machen, weil es die gleichzeitige Bestimmung von Metall und Schwefel in einer Probe in allen Salzen erlaubt, deren Metall in Form von Alkalikarbonat gefällt werden kann. Wird das Filtrat dieser Fällung mit etwas Kali- oder Natronlauge gekocht, so entweicht alles Pulegon. Durch Hinzufügen von etwas freiem Brom kann man das in Lösung befindliche Alkalisulfit zu Sulfat oxydieren und die Schwefelsäure in bekannter Weise bestimmen. Dagegen erhält man — wie besonders erwähnt zu werden verdient — keine Abscheidung von Baryumsulfat, wenn man eine wässrige Lösung des Barytsalzes der Sulfosäure mit Brom erwärmt, was für die feste Bindung des Schwefligsäureradikals spricht. Ferner ist zu bemerken, daß alle Salze der Permanganatprobe gegenüber sich wie gesättigte Verbindungen verhalten. Die Pulegonsulfosauren Salze entsprechen also keineswegs den Salzen, die man durch Addition von Natriumbisulfit an gesättigte Ketone erhält[1]), vielmehr ist anzunehmen, daß sich SO_3H_2 an die semicyclische Bindung anlagert (s. Abschn. II). Die Analysen der Salze haben nun ergeben, daß sie nicht der einfachen Formel $C_{10}H_{17}O . SO_3M$ entsprechen, sondern, daß sie alle Wasser in recht fester Bindung enthalten, denn die schön kristallisierten Salze zeigen keinerlei Neigung zur Verwitterung und verlieren auch bei 100° nur langsam an Gewicht und weniger als ihrem durch die Elementaranalyse ermittelten Wassergehalt entsprechen würde, obgleich sich bei jener Temperatur schon Zersetzungserscheinungen (Pulegongeruch) bemerklich machen. Wahrscheinlich beteiligt sich also die Karbonylgruppe an der Wasseraufnahme. Von den analytischen Resultaten seien folgende mitgeteilt:

Bariumsalz:

ber. für	$(C_{10}H_{17}O_4S)_2Ba$	$+ 2H_2O$	$+ 3H_2O$	gefunden
Ba	22.75	21.47	20.87	20.6; 20.9; 20.95; 21.2
S		10.02	9.75	6.69

1) Baeyer hat (Ber. 28, 652 [1895]), um den Ketoncharakter des Pulegons zu erweisen, die Verbindung mit Natriumbisulfit geschüttelt und nach längerer Einwirkung bei Gegenwart von Alkohol die Abscheidung eines festen Natriumsalzes erzielt, das, wie besondere Versuche gezeigt haben, identisch ist mit dem Salz, das man durch Umsetzung der oben beschriebenen Salze der Pulegonsulfosäure mit Natriumkarbonat erhält.

Calciumsalz:

	ber. für $(C_{10}H_{17}O_4S)_2Ca$	$+ 2H_2O$	$+ 3H_2O$	gefunden
Ca	7.91	7.37	7.14	7.28; 7.30; 7.18
S	12.63	11.80	11.42	11.43; 11.57;

Magnesiumsalz:

	ber. für $(C_{10}H_{17}O_4S)_2Mg$	$+ 2H_2O$	$+ 3H_2O$	
S	13.04	12.15	11.75	12.16

Bleisalz:

	ber. für $(C_{10}H_{17}O_4S)_2Pb$	$+ 2H_2O$	$+ 3H_2O$	
Pb	30.80	29.20	28.46	28.30

Zinksalz:

	ber. für $(C_{10}H_{17}O_4S)_2Zn$	$+ 2H_2O$	$+ 3H_2O$	
Zn	12.29	11.52	11.16	11.06; 11.17.

Nach Versuchen, für die ich Herrn Prof. Heubner zu Dank verbunden bin, sind die Salze, sofern sie ungiftige Melalle enthalten, physiologisch ganz unwirksam.

II. Das Verhalten von Sabinol und von Sabinen gegen schweflige Säure.

Die Beobachtungen über die Additionsfähigkeit von schwefliger Säure an Pulegon und Campherphoron ließen die Frage entstehen, ob es die tertiär-semicyclische Kohlenstoffbindung allein ist, die hier die Additionsfähigkeit von SO_3H_2 begünstigt, ob semicyclische Bindungen überhaupt dazu befähigt sind, oder ob in den genannten Fällen die Nahstellung von Carbonyl (bezw. anderer sauerstoffhaltiger Gruppen) zur Doppelbindung der Kohlenstoffatome für die Aufnahmefähigkeit bestimmend ist. Um die erste Frage zu entscheiden, hätte es am nächsten gelegen, den dem Pulegon zu Grunde liegenden Kohlenwasserstoff (III) auf sein Verhalten gegen schweflige Säure zu untersuchen. Diese Verbindung konnte aber unter den gegenwärtigen Verhältnissen nicht beschafft werden. Versuche die u. a. mit der Pulegensäure (IV), dem β-Pinen (V) und einer Reihe anderer geeigneter Verbindungen angestellt wurden, verliefen bisher ergebnislos.

Dagegen wurde bei Anwendung von S a b i n o l (VI)

III. IV. V. VI.

$C(CH_3)_2$ —COOH =$C(CH_3)_2$ CH_2 CH_2 —OH C_3H_7

ein positives Resultat erhalten.

Läßt man Sabinol, das durch Destillation mit Wasserdampf gereinigt und dann im Vakuum fraktioniert wurde ($Kp._{16}$ 95—97°), über wässriger concentrierter schwefeliger Säure stehen, so verwandelt es sich in eine kautschukartige Masse, die ein Gemenge verschiedener Körper vorstellt. Viel vollständiger und besser vollzieht sich die Reaktion, wenn man das Sabinol mit der schwefeligen Säure auf der Schüttelmaschine einige Stunden durchschüttelt. Es scheidet sich dann eine mehr körnige weiße Substanz aus, von der man die wässrige Flüssigkeit abgießen kann. Bei deren Verdunstung in einer flachen Schale scheidet sich in guten Krystallen eine Sulfosäure (A) aus. Die von der ursprünglichen sauren Flüssigkeit befreite feste weiße Masse wird mit Ammoniak durchgerieben und die ammonikalische Lösung mit Aether ausgezogen. Dieser nimmt noch anhaftendes Sabinol und etwas von einem n e u t r a l e n Körper (B) auf. Aus der ammoniakalischen Lösung fällt Schwefelsäure noch mehr von der neuen Sulfosäure A. Der von Ammoniak nicht gelöste Rückstand ist in allen üblichen Lösungsmitteln völlig unlöslich. Er wird durch Erwärmen mit Alkohol auf dem Wasserbad ausgekocht, filtriert und mit Alkohol und Aether ausgewaschen. Das rückständige Produkt (C) ist amorph und schmilzt bei etwa 213—215°.

Statt die Umwandlung des Sabinols mit w ä s s r i g e r schwefeliger Säure vorzunehmen, kann man eine Lösung desselben in etwa 3 Teilen absolutem A l k o h o l mit SO_2-Gas sättigen. Wenn man dabei mit Eis kühlt, bleibt die Flüssigkeit klar und setzt erst nach einiger Zeit etwas von der weißen unlöslichen Substanz (C) ab. Die Menge, die man von dieser letzeren erhält, hängt auch sonst von den Versuchsbedingungen ab. Manchmal ist sie sehr gering, sie vermehrt sich aber erheblich, wenn die Flüssigkeit beim Einleiten der schwefeligen Säure sich erwärmt oder vorher absichtlich erwärmt wird.

Wenn man die Sulfosäure (A) gewinnen will, sättigt man demnach eine alkoholische Lösung von Sabinol bei 0° mit SO_2, läßt die Flüssigkeit einen Tag stehen, gießt dann (ev. nach dem Abfiltrieren der ausgeschiedenen Substanz C) in eine Schale aus und läßt Alkohol und schwefelige Säure abdunsten. Die entstandene Säure scheidet sich schließlich krystallinisch aus, enthält aber noch etwas ölige Substanz und von dem neutralen krystallinischen Körper B. Um die Säure von diesen Bestandteilen zu befreien, wird mit Ammoniak aufgenommen und die ammoniakalische Lösung mit Aether ausgeschüttelt. Nach Entfernung des Aethers hinterbleibt B, meist von öligen Anteilen (sabinolhaltig) durchtränkt. Man saugt letztere auf Ton ab und krystallisiert B aus Alkohol oder auch Aceton um. In reinem Zustand bildet dieser neutrale Körper (B) wollige Nadeln vom Schmelzpunkt 147—148°.

Das Hauptreaktionsprodukt zwischen Sabinol und schwefeliger Säure ist die gut und leicht krystallisiert zu erhaltende Sulfosäure A. Man kann sie aus ammoniakalischer Lösung durch Neutralisation mit Schwefelsäure abscheiden. Dann fällt sie als bald erstarrendes Öl. Noch besser erhält man sie, wenn man in die frisch bereitete, farblose und nicht zu verdünnte ammoniakalische Lösung unter guter Eiskühlung Salzsäuregas einleitet. Die Säure fällt dann sofort fest und rein weiß aus, wird abfiltriert, um anhaftenden Salmiak zu entfernen, mit wenig Wasser nachgewaschen und schließlich aus wenig verdünntem Methylalkohol umkrystallisiert. Bei langsamem Verdunsten einer wässrigen Lösung kann man sie in großen Krystallen erhalten. In Alkohol ist sie sehr leicht löslich, in kaltem Wasser viel schwerer. In reinem Zustand schmilzt die Säure bei 98—99°. Bei längerem Liegen an der Luft färbt sie sich erst gelb, dann rot und zerfließt allmählich zu einer schmierigen Masse. Die Lösungen sind rechtsdrehend.

Das Ammoniaksalz ist krystallinisch und im Gegensatz zu der freien Säure gut haltbar. Man kann es durch Abdunsten einer ammoniakalischen Lösung der Säure herstellen. In schneeweißem Zustande fällt es aus, wenn man eine Lösung der Säure in Aether oder in wasserfreiem Alkohol mit Ammoniakgas sättigt.

Analyse:	Berechnet für $C_{10}H_{15}SO_3NH_4$	Gefunden:
C	51.43	51.41
H	8.21	8.28

Das Calciumsalz wurde durch Neutralisation einer wässerigen Lösung der Säure mit Calciumcarbonat hergestellt. Es ist sehr löslich in Wasser und kleinkrystallinisch.

Analyse: Berechnet für $(C_{10}H_{15}SO_3)_2Ca$		Gefunden:	
Ca	8.21	8.36	8.37

Die Analyse der freien bei 98—99° schmelzenden Säure ergab:

Berechnet für $C_{10}H_{15}SO_3H$		Gefunden:			
C	55.5	55.39	55.32	55.71	—
H	7.4	7.65	7.63	7.73	—
S	14.8	—	—	—	14.01

Es unterliegt also keinem Zweifel, daß man es mit einer Sulfosäure der Formel $C_{10}H_{15}SO_3H$ zu tun hat, deren Bildung durch Addition von SO_3H_2 an Sabinol $C_{10}H_{15}OH$ unter gleichzeitigem Austritt von 1 Mol. Wasser erfolgt ist.

Die Säure verhält sich gegen Permanganat stark ungesättigt, ebenso gegen Brom. Durch Addition von 1 Mol. Brom an die Säure (in ätherischer oder wässriger Lösung) erhält man eine amorphe Substanz.

Welche Konstitution kommt nun der Sulfosäure $C_{10}H_{15}SO_3H$ zu?

Die Richtigkeit der für Sabinol jezt allgemein angenommenen Formel (VI) vorausgesetzt, kann dessen leichte Additionsfähigkeit zu SO_3H_2 durch das Vorhandensein der Methen-bindung oder die leichte Sprengbarkeit des Dreirings bedingt sein. Es war daher angezeigt zu ermitteln, wie sich Sabinol nach Aufhebung der semicyclischen Doppelbindung, also reduziertes Sabinol, d. h. Thujylalkohol (s. Abschnitt III) gegen schwefelige Säure verhalten würde. Dabei zeigte es sich, daß bei Aufhebung der Methengruppe die Additionsfähigkeit zu SO_3H_2 ganz verschwindet. Allerdings wäre es voreilig nun schon zu schließen, daß bei der Einwirkung von schwefeliger Säure auf Sabinol selbst der Dreiring intakt bleibt, denn ich habe wiederholt auf die Tatsache aufmerksam gemacht, daß die Festigkeit des Dreirings in starker Abhängigkeit von der sonstigen Konfiguration des Molekuls steht, daß sie z. B. in der α-Thujaketosäure außerordentlich viel geringer ist, als in der aus dieser durch Oxydation entstehenden, zugehörigen Dicarbonsäure [1]).

Auch wenn die Addition an der Methengruppe einsetzt, könnte doch eine gleichzeitige Auflösung des tricyclischen Rings unter Entstehung einer Aethylenbindung erfolgen (analog dem Übergang von Thujon in Carvotanaceton oder in Isothujon). In diesem Fall könnte man es mit einer Dihydrocymolsulfosäure, etwa der

1) Wallach, Terpene und Campher, Aufl. II, S. 503.

Formel VII zu tun haben. Eine solche müßte aber, allen Erfahrungen nach, zur Abspaltung von SO_3H_2 unter Bildung von Cymol neigen. Das trifft jedoch nicht zu. Weder beim Erhitzen der bei 98—99° schmelzenden Sulfosäure mit Dimethylanilin, noch mit Alkali tritt Cymol in merklicher Menge auf. Ebensowenig beim trockenen Erhitzen ihrer Salze, wobei sich Entbindung von Schwefelwasserstoff bemerklich macht.

Daß der Dreiring in der Sulfosäure nun tatsächlich noch erhalten ist, geht aus Folgendem hervor:

Wenn man die Säure $C_{10}H_{15}SO_3H$ in einem Kupfergefäß einige Stunden mit Kalilauge (1:1) kocht, so entsteht, unter Abspaltung von Kaliumsulfit und reichlicher Bildung harzartiger Produkte (teilweise von saurem Charakter) ein flüchtiges, pfefferminzartig riechendes Öl, das mit Wasserdampf aus der Reaktionsmasse abgetrieben werden kann. Die Ausbeute ist gering. Da das Öl sich Permanganat gegenüber ungesättigt verhielt, so wurde das Wasserdampfdestillat bei Gegenwart von Chlorpalladium mit Wasserstoff behandelt, bis die sich langsam vollziehende Wasserstoffaufnahme beendet war. Es wurde so ein Präparat von folgenden Eigenschaften erhalten: Siedepunkt 199—204°, $d_{19} = 0.9165$, $n_D = 1.4541$. Diese Substanz, in der ein Alkohol zu vermuten war, wurde in Eisessiglösung mit Chromsäure (ber. auf 1 Atom O) oxydiert. Dabei entstand ein Keton, das leicht als Thujon identifiziert werden konnte [1]).

Das reine Keton siedete bei 199—200°.

Analyse:	Berechnet für $C_{10}H_{16}O$	gefunden:	
C	78.88	78.79	78.65
H	10.60	10.36	10.60

Das Oxim blieb flüsssig. Bei der Oxydation mit Permanganat lieferte das Keton eine Ketosäure. Semicarbazon Schmelzpunkt 182—183° unter Aufschäumen. Die daraus freigemachte Säure wurde in Berührung mit α-Thujaketosäure sofort fest.

Unerwarteter Weise gab die Analyse der durch Behandlung des ursprünglichen Reaktionsproduktes mit Wasserstoff erhaltenen, bei 199—204° siedenden Substanz Werte, die gleichfalls auf $C_{10}H_{16}O$ (nicht auf $C_{10}H_{18}O$) stimmen. Das direkte Umsetzungsprodukt der Sulfosäure $C_{10}H_{15}SO_3H$ mit Kali enthielt also durch freien Wasser-

1) Die Isolierung erfolgte über das schlecht charakterisierte, in Alkohol und in Äther leicht lösliche Semicarbazon in derselben Weise und unter Eintreten derselben Erscheinungen, wie es im folgenden Abschnitt III gelegentlich der Reindarstellung von Thujon beschrieben werden wird.

stoff reduzierbare Verunreinigungen, oder schon Thujon[1]) neben dessen Enolform.

Diese Entstehung von Thujon IX aus der Sulfosäure erklärt sich leicht, wenn man für die Säure die Formel VIII annimmt, deren Bildung bei der Umsetzung von Sabinol mit schwefeliger Säure sich auch deuten ließe.

VII. CH_3 — SO_3H — C_3H_7

VIII. CH_3 — SO_3H — C_3H_7 → CH_3 — OH — C_3H_7 → XI. CH_3 — =O — C_3H_7

Gegen die Annahme der Formel VIII spricht nur, daß eine Säure dieser Konstitution bei vorsichtiger Oxydation mit Permanganat α-Thujaketosäure liefern sollte, was nach den bisherigen Beobachtungen nicht zutrifft. Daher kann die Konstitution der Sulfosäure noch nicht als einwandfrei festgestellt gelten.

Von den gleichzeitig mit der Sulfosäure $C_{10}H_{15}SO_3H$ bei der Einwirkung von schwefeliger Säure auf Sabinol sich bildenden Nebenprodukten, entsteht der neutrale, krystallisierte Körper (B) vom Schmelzpunkt 147—148° stets nur in geringer Menge und es muß dahingestellt bleiben, ob er nicht von einer das Sabinol begleitenden Fremdsubstanz stammt.

Die ganz reine Verbindung ist in Alkohol und Äther nicht leicht löslich und wird, wie schon oben bemerkt, aus heißem Alkohol oder aus Aceton krystallisiert, in wolligen Nadeln erhalten.

Analyse:	C	65.08	65.36 %	S 8.06 %
	H	7.64	7.86 %	

Annähernd stimmen diese Werte auf $2C_{10}H_{16}O + 1SO_2$. Die Verbindung ist ganz gesättigt. Falls sie aus dem Sabinol stammt, hat man es wohl also mit einem Polymerisationsprodukt zu tun, in dem Athylenbindungen nicht mehr vorhanden sind.

Die amorphe Verbindung (C) vom Schmelzpunkt 213—215° ist wegen ihrer Unlöslichkeit nicht umkrystallisierbar und daher schwer in einwandfrei reinem Zustand zu erhalten. Die bisher ausgeführten Analysen ergaben folgende Durchschnittszahlen:

1) Die oben mitgeteilten Konstanten deuten auch auf Thujon-Gehalt hin.

$$C = 53{,}72\ \%; \quad H = 7{,}0\ \%.$$

Die Formel $C_{20}H_{34}O_7S_2$ ($2C_{10}H_{16}O + 2SO_2 + H_2O$) verlangt $C = 53.29$; $H = 7.61\ \%$.

Daß die Verbindung esterartiger Natur ist, geht daraus hervor, daß sie beim Kochen mit Natriumalkoholat in Lösung übergeht, wobei sich eine Säure von harzartiger Beschaffenheit bildet.

Sehr wahrscheinlich liegt auch hier eine polymere Verbindung vor, deren Entstehung wegen ihrer Unlöslichkeit für die Auffindung von Sabinol in Gemischen wird von Nutzen sein können.

Überhaupt wird die große Reaktionsfähigkeit des Sabinols gegenüber der schwefeligen Säure unter Bildung einerseits einer gut krystallisierenden und leicht zu isolierenden Sulfosäure, andererseits (unter geeigneten Bedingungen) der durch Unlöslichkeit ausgezeichneten Substanz, es gestatten, das Sabinol, das in ätherischen Ölen vermutlich häufiger sich findet, als man bisher weiß, leicht nachzuweisen.

Auch wird sich die schweflige Säure als Reagens bei der Untersuchung ätherischer Öle voraussichtlich noch ausgiebiger verwerten lassen. So hat sich z. B. herausgestellt, daß auch käufliches Sabinen (über Natrium rektifiziert und ganz frei von Sabinol) einen schwerlöslichen weißen Körper abscheidet, wenn man in die alkoholische Lösung des Kohlenwasserstoffs SO_2 einleitet. Diese Verbindung entsteht indeß nur in geringer Menge (stammt also vielleicht aus einem beigemengten besonderen Kohlenwasserstoff). Sie ist amorph und schwefelhaltig, verkohlt vor dem Schmelzen und ist ganz verschieden von der aus Sabinol entstehenden unlöslichen, bei 213—215° schmelzenden Substanz.

III. Über Thujon.

1) Wie ich in einer ausführlichen Untersuchung früher nachgewiesen habe[1]), kommt das Thujon (Tanaceton) in mehreren stereoisomeren Modifikationen vor. Eine linksdrehende (α) ist (neben Fenchon) wesentlich in Thujaöl enthalten, eine rechtsdrehende (β) bildet einen Hauptbestandteil namentlich des Tanacetöl und des Wermutöls. Die rechtsdrehende (β-) Modifikation ist die beständigere. Durch Alkali läßt sich α-Thujon in β-Thujon überführen. Semmler[2]) hat gefunden, daß, wenn man Sabinol in amylalkoholischer Lösung mit Natrium reduziert, ein Alkohol ent-

---

1) Ann. d. Ch. **336**, 247 (1904).
2) Ber. **33**, 1461 (1900).

steht, der bei der Oxydation β-Thujon liefert. Angaben über das optische Verhalten dieses Präparats hat Semmler allerdings nicht gemacht, aber daraus ein krystallisiertes Oxim vom Schmelzpunkt 52° erhalten, während Links-Thujon nur ein flüssiges Oxim gibt. Da nun Alkali invertierend wirken kann, schien es mir nicht ausgeschlossen, daß man vom Sabinol aus auch zu einem linksdrehenden Thujon würde kommen können, wenn man die Reduktion in anderer Weise bewerkstelligt. Der Versuch hat das bestätigt. Wird Sabinol bei Gegenwart von Chlorpalladium mit freiem Wasserstoff reduziert, so entsteht kein β-Thujylalkohol (Tanacetylalkohol), sondern — neben Kohlenwasserstoff — ein linksdrehender sekundärer Alkohol, der bei vorsichtiger Oxydation wesentlich Links-Thujon liefert. Die Reduktion des Sabinols mit Chlorpalladium wurde in wäßriger Suspension unter Hinzufügen von wenig Methylalkohol vorgenommen. Sobald die Wasserstoffaufnahme merklich nachließ, wurde mit Wasserdampf abgeblasen, das Destillat neuerdings der Reduktion unterworfen, der Vorgang noch einmal wiederholt und schließlich wurden noch vorhandene ungesättigte Anteile durch Oxydation mit 1%iger Permanganatlösung bei niederer Temperatur fortgenommen.

Das so gewonnene gesättigte Reaktionsprodukt erwies sich als nicht einheitlich. Es sott in ziemlich weiten Grenzen. Der niedrig siedende Anteil bestand aus Kohlenwasserstoff, aus dem die letzten Anteile vorhandenen, höher siedenden Alkohols durch Rektifikation über metallischem Natrium fortgenommen werden konnten. Der gesättigte Kohlenwasserstoff zeigte folgende Eigenschaften:

$$\text{Kp. } 161—163^{\circ}, \quad d_{21} = 0.795, \quad n_D = 1.4375.$$

Der Katalysator hatte also zum Teil Wasser abspaltend gewirkt, was bei einem sekundären Alkohol von der dem Sabinol (bezw. Dihydrosabinol) zugeschriebenen Formel auffallend ist. Die Hauptmenge des Reduktionsprodukts bestand aber aus einem gesättigten Alkohol, dessen Hauptmenge folgende Konstanten aufwies:

$$\text{Kp. } 206—208^{\circ}, \quad d_{20} = 0.9090, \quad n_D = 1.4576, \quad M = 46.19.$$

Der Alkohol ist linksdrehend. In wässriger Suspension wird er durch Chromsäure (bei Gegenwart von etwas Schwefelsäure) nur schwer angegriffen, dagegen schnell unter spontaner Erwärmung oxydiert, wenn man zu einer Eisessiglösung die berechnete Menge von in wenig Wasser gelöstem Chromtrioxyd hinzufügt. Das mit Wasserdampf abdestillierte Oxydationsprodukt wurde abgehoben, zu einer konzentrierten Lösung von essigsaurem Semicarbazid ge-

setzt und so viel Alkohol hinzugefügt, daß gerade eine klare Lösung entstand. Es bildet sich schnell ein Semicarbazon, dessen Herausarbeitung jedoch durch seine ungünstige Beschaffenheit (große Löslichkeit in Alkohol, Äther und beigemengten Verunreinigungen) erschwert wird. Wenn man die Hauptmenge des Alkohols in einer offenen Schale verdunsten läßt, scheidet sich eine von Öl durchtränkte Krystallmasse aus. Nun wäscht man durch Hinzufügen von Wasser die unorganischen Salze fort und reibt nach Abgießen des Wassers den Rückstand mit Sodalösung durch, um freie Säure abzustumpfen [1]). Durch die mit Hülfe von etwas Alkohol in einen Rundkolben gespülte Masse wird zur Entfernung aller flüchtiger Bestandteile Wasserdampf geleitet. Im Kolbenrückstand bleibt das Semicarbazon, das, in der Wärme flüssig, beim Abkühlen erstarrt. Man gießt nun die darüberstehende Flüssigkeit ab [1]), wäscht mit Wasser nach, fügt sodann überschüssige feste Oxalsäure in den Kolben und destilliert abermals mit Wasserdampf. Dabei spaltet sich das vorhandene Semicarbazon sehr leicht und ein farbloses Keton geht über.

Dies Keton zeigt folgende Eigenschaften:

$$\text{Kp. } 203-205^0, \quad d_{20} = 0.907, \quad n_D = 1.4507.$$

Es ist linksdrehend. S = 1.1666, L (Methylalkohol) = 0.60805, $p = 16{,}1\,^0/_0$, $d = 0.8115$, $t = 18^0$, $\alpha = -4^0$, $[\alpha]_D = -30{,}62$. Das Präparat war leicht als Thujon zu identifizieren, denn es liefert bei der Oxydation mit Permanganat α-Thujaketosäure. Seine Linksdrehung übertrifft erheblich die früher für ein aus Thujaöl dargestelltes Präparat gefundene $(-10{,}23^0)$ [2]). Es muß aber dahingestellt bleiben, ob eine dritte Modifikation des Thujon vorliegt, deren Existenz früher schon angedeutet, aber nicht sicher erwiesen ist [3]). Wegen der leichten Invertierbarkeit des Thujons ist die Entscheidung schwierig [4]). Das vorliegende Präparat gab ein

---

1) Die wässrige Flüssigkeit, die schwach alkalisch sein muß, war gelb gefärbt und enthielt eine mit Säuren fällbare gelbe organische Substanz, die das feste Semicarbazon verunreinigt und auch durch Ammoniak fortgewaschen werden kann.

2) Ann. d. Ch. **336**, 263 (1904).

3) Ebenda **336**, 258.

4) Die Isolierung des (neben λ-Fenchon) im Thujaöl enthaltenen α-Thujon wird man auf Grund der neueren Erfahrungen gegen früher verbessern können, indem man ähnlich verfährt, wie oben beschrieben, d. h. Thujaöl kurze Zeit mit Semicarbazid stehen läßt und dann das unaugegriffene Fenchon durch Wasserdampf entfernt (vergl. auch Ann. **353**, 215 (1907) und das rückständige Semicarbazon mit Oxalsäure zerlegt. Vielleicht läßt sich so auch aus Thujaöl ein stärker drehendes Thujon gewinnen.

flüssiges Oxim, das nach längerem Stehen Ansätze zur Krystallisation zeigte, also wohl auch geringe Mengen β-Thujon enthielt. Das Semicarbazon des stark linksdrehenden Präparats bildete sich sehr schnell, ist aber in Methylalkohol und Äther z. T. recht löslich und schlecht krystallisierbar. U. a. wurden nach häufigem Umkrystallisieren gut ausgebildete Krystalle vom Schmelzpunkt 153—154° isoliert, die wohl Mischkrystalle vorstellen, deren Entstehen ich gelegentlich einer früheren ausführlichen Untersuchung über die Semicarbazone des Thujon auch beobachtet hatte[1]). Zur sicheren Charakterisierung des λ-Thujons ist das Semicarbazon jedenfalls nicht besonders geeignet. Um die Frage der Identität des aus Sabinol gewonnenen Links-Thujon mit dem im Thujaöl enthaltenen entscheiden zu können, sind erneute Versuche unter Bedingungen nötig, die eine Umwandlung von α- in β-Thujon möglichst verhindern. Ein sorgfältiger Vergleich von verschiedenen Modifikationen des Thujon ist um so erwünschter als sie kleine Abweichungen betreffs der Siedepunkte aufzuweisen scheinen, die festzustellen wären.

2) Gelegentlich anderer Versuche wurde beobachtet, daß Thujon-haltige Öle, die längere Zeit aufbewahrt waren, einen mehr oder weniger bedeutenden Säuregehalt aufweisen. Eine nähere Untersuchung zeigte, daß diese Säure in der Hauptmenge aus α-Thujaketosäure $C_{10}H_{16}O_3$ besteht, der eine kleinere Menge einer harzartigen, nicht an Semicarbazid gehenden, Säure beigemengt ist. Dieser Säuregehalt ließ sich ebenso wie in Rohölen, auch in Anteilen Thujonhaltiger Öle vom Siedepunkt des Thujons, die, zwecks Isolierung dieses Ketons, vor einer längeren Reihe von Jahren herausfraktioniert worden waren, feststellen.

Was den Nachweis der α-Thujaketosäure in den betreffenden Ölen anbelangt, so ist er sehr einfach. Man schüttelt die zu prüfende Flüssigkeit mit Natronlauge durch, entzieht der abgehobenen alkalischen Lösung anhaftende neutrale Verunreinigungen durch Äther, säuert mit Schwefelsäure an und äthert die freigemachte Säure aus. Nach Entfernung des Äthers hinterbleibt ein meist braun gefärbter, nur schwer krystallisierender Syrup, den man mit wenig Alkohol aufnimmt, zu einer konzentrierten Lösung von essigsaurem Semicarbazid setzt, unter Hinzufügen von so viel Alkohol, daß gerade eine klare Lösung entsteht. Es scheidet sich, meist schon nach kurzer Zeit, ein schwerlösliches Semicarbazon aus, das man abfiltriert und durch Nachwaschen mit etwas Alkohol

1) Ann. 336, 255, 269, 270 (1904).

und Äther von färbenden Verunreinigungen befreit. Man kann das Semicarbazon zwecks weiterer Reinigung in Ammoniak lösen und nach dem Filtrieren mit Essigsäure fällen. Zerlegt man nun das weiße Semicarbazon durch schwaches Erwärmen mit Schwefelsäure (1 : 1), so nimmt Äther die freigewordene Säure auf, die nach Verdunsten des Lösungsmittels sehr bald erstarrt, farblos ist und den richtigen Schmelzpunkt aufweist. In den von dem Roh-Semicarbazid abfiltrierten Laugen findet sich die harzartige Säure. Die Anwesenheit von α-Thujaketosäure wurde in Präparaten, die seit mehreren Jahren im hiesigen Laboratorium aufbewahrt wurden, u. a. im Salbeiöl und in wesentlich aus Thujon bestehenden Fraktionen des Thujaöls, Tanacetöls, Wermutöls, Artemisiaöls gefunden. Die Firma Schimmel & Co. in Miltitz hatte die Güte auf meine Bitte hin festzustellen, daß in dort lagerndem Thujon und Tanacetöl gleichfalls ansehnliche Mengen α-Thujaketosäure enthalten waren. Z. B. zeigte Rainfarnöl vom Lager (geringe Mengen in großer Flasche) eine Säurezahl von 4,6 und durch Ausschütteln mit Natronlauge wurden aus 0,6 kg 14 gr Säure erhalten.

Hier aufbewahrtes, vor längerer Zeit von Schimmels bezogenes „Tanaceton“ enthielt etwa 11 Vol. % durch Kalilauge (1 : 2) ausschüttelbarer saurer, zumeist aus α-Thujaketosäure bestehender Bestandteile. Eine dreimal destillierte, jetzt ganz dunkel gefärbte Fraktion aus Thujaöl vom s. Z. vermerkten Siedepunkt 200—212°, die sich in großer, nur zu kleinem Teil gefüllter Flasche befand, gab nicht weniger als 20 Vol. % an Kalilauge ab. Sehr reich an Säure erwies sich ferner eine farblose Fraktion (vermerkter Siedep. 200—205°) aus Wermutöl, das von früheren Versuchen[1] (1904) her in zwar nicht ganz voller aber gut verkorkter kleiner Flasche gestanden hatte. Aus 50 ccm dieses Öls ließen sich aus den mittelst Alkali daraus entfernten sauren Bestandteilen 9,5 gr rohes Semicarbazon von α-Thujaketosäure isolieren, was etwa 7 gr. an der Säure entspricht.

Dieser so überraschende und bedeutende Gehalt an α-Thujaketosäure in mehrfach (teils mit Wasserdampf, teils unter gewöhnlichem Druck) destillierter Fraktionen, deren Siedepunkt schon etwas oberhalb des Thujonsiedepunktes (200—201°) lag, erklärt sich nun daraus, daß, wie sich jetzt herausgestellt hat, die α-Thujaketosäure mit Wasserdampf merklich und mit den Dämpfen siedenden Thujons nicht unerheblich flüchtig ist. Enthält ein

1) Ann. **336**, 268 (1904).

Thujon-Präparat also schon von der Säure, so reichert sich diese bei fortgesetztem Fraktionieren in den höher siedenden Anteilen an. Was noch zu erklären bleibt, ist, wie die bekanntlich durch Ringsprengung aus dem Thujon, $C_{10}H_{16}O$, durch Oxydation erhältliche Ketonsäure $C_{10}H_{16}O_3$ in den betreffenden Präparaten, die mit chemischen Oxydationsmitteln nicht in Berührung kamen, ursprünglich überhaupt entstanden ist?

Von einem gesättigten Keton sollte man eine Neigung zur Sauerstoffaufnahme nicht erwarten. Man darf indes annehmen, daß Thujon zur Enolisierung neigt und als ungesättigte Verbindung reagiert. Darauf beruht wohl auch seine auffallend leichte Angreifbarkeit durch Permanganat.

Es ließ sich nun feststellen, daß Thujon auch durch Wasserstoffsuperoxyd, das sich ja beim Verdunsten mancher Bestandteile ätherischer Öle gern bildet, aufspaltbar ist, wenn man das Keton einige Wochen bei Gegenwart von etwas Natronlauge mit starken Perhydrol-Lösungen auf der Schüttelmaschine durchschüttelt. Immerhin verläuft der Prozeß recht träge und kaum merklich bei Abwesenheit von Alkali. Auch findet keine wahrnehmbare Sauerstoffaufnahme statt, wenn man durch säurefreies, feuchtes Thujon einen Strom Sauerstoffgas hindurchleitet. Man könnte daran denken, daß die Sauerstoffaufnahme durch das Licht beschleunigt wird und daß dabei ein intramolekularer Prozeß sich abspielt. Daß sie auch im Dunklen stattfinden kann, geht aus der erwähnten Mitteilung von Schimmels hervor, der zu Folge aus 4 Kg. Thujon, die vor 3 Jahren über Alkali[1]) mit Wasserdampf und dann im Vakuum destilliert waren und seitdem in einer 5 kg-Flasche im Dunklen im Keller gestanden hatten, jetzt 9 gr Säure isoliert werden konnten. Eine ganz befriedigende Erklärung für das Vorhandensein von Ketonsäure in unseren Präparaten kann auf experimenteller Grundlage noch nicht gegeben werden. Die beobachteten Tatsachen dürften aber schon nutzbar zu machen sein, um in längere Zeit aufbewahrten ätherischen Ölen *einen Thujon-Gehalt nachzuweisen.*

---

1) Es muß bei der Gelegenheit darauf aufmerksam gemacht werden, daß konzentriertes Alkali Thujon stark angreift. Erwärmt man Thujon z. B. mit etwas festem Ätznatron, so wird es bei 100° und noch schneller und weitgehender bei höherer Temperatur unter Braunfärbung verändert. Dabei entstehen stark verharzte Produkte von saurer Beschaffenheit, unter denen Carvacrol, nicht aber α-Thujaketosäure aufgefunden werden konnte.

Man braucht die Präparate nur mit Natronlauge auszuschütteln und in der oben angegebenen Weise die vorhandene Säure auf einen Gehalt an α-Thujaketosäure zu prüfen.

An den im Vorstehenden kurz mitgeteilten Untersuchungen haben sich die Herren Dr. W. Reggelin und Dr. F. Riesener in dankenswerter Weise beteiligt.

Alte und junge Saumtiefen.

Von

Hans Stille.

Vorgelegt in der Sitzung vom 18. Juli 1919.

Einleitung. Das retrospektive Verfahren zur Deutung gegenwärtiger morphologischer Erscheinungen.

Seitdem Charles Lyell und sein deutscher Vorläufer K. E. A. v. Hoff in den ersten Jahrzehnten des vorigen Jahrhunderts gezeigt haben, daß die heute auf unserer Erde wirksamen Kräfte ausreichen, um die Vorgänge der Vorzeit zu erklären, bedient man sich zur Deutung der Verhältnisse der Vorzeit der für uns jetzt selbstverständlichen „aktualistischen“ Methode, d. h. also der vergleichenden Heranziehung der Verhältnisse der Jetztzeit. Als „ontologische“ Methode hat Joh. Walther[1]) sie bezeichnet und

1) Joh. Walther, Einleitung in die Geologie als historische Wissenschaft. Jena, 1893/94.

besonders auf die bionomischen Probleme der Vorzeit in Anwendung gebracht.

Umgekehrt können nun mancherlei Erscheinungen der Gegenwart durch die vergleichende Untersuchung entsprechender Verhältnisse der Vorzeit ihrer Deutung nähergebracht werden. Das Einzelbild des gegenwärtigen Zustandes ermöglicht in vielen Fällen nicht die ausreichende Erklärung seines Zustandekommens; dagegen ist der Geologe bei den gleichen Erscheinungen der Vorzeit oft in der glücklichen Lage, die vorangegangenen Stadien enthüllen und dadurch den Werdegang der Erscheinungen erklären zu können. Daraus ergibt sich dann die Nutzanwendung auf die Erscheinungen der Jetztzeit.

Während die aktualistische Methode die Gegenwart zur Erklärung der Vergangenheit benutzt, schließen wir also bei dieser „retrospektiven“ Forschungsart aus der Vergangenheit auf die Gegenwart.

Das heutige Meeresbecken sagt uns z. B. wenig über seine Entstehung, ja verführt sogar leicht zu falschen Schlüssen, indem wir etwaige in seiner Peripherie aufsetzende Verwerfungen ohne weiteres mit Senkungserscheinungen des Beckens in Verbindung bringen. Auf sichererem Boden stehen wir aber hinsichtlich der Meere der Vorzeit in solchen Gebieten, die heute landfest sind und dabei eine genauere Untersuchung der Untergrundsverhältnisse, vor allem auch der zeitlichen Folge der Ereignisse, die mit der Bildung des alten Meeresraumes in Zusammenhang gestanden haben, zulassen. Und daß wir auf diesem Wege zu einem abweichenden Ergebnisse kommen gegenüber der ziemlich allgemein hinsichtlich der Entstehung unserer gegenwärtigen Meere geltenden Auffassung, nämlich zu dem, daß die Meeresräume nicht eingebrochen sind, sondern sich langsam und bruchlos eingesenkt haben, habe ich an anderer Stelle zu zeigen versucht, — und daraus ergeben sich die Konsequenzen für die Deutung der gegenwärtigen Meere. Das ist ein solcher Fall der retrospektiven Behandlung eines Problems der Gegenwart. Auch sonst ist diese Methode geübt worden; ich erinnere an die Mittelschwelle des Atlantik, deren Natur E. Haug durch Vergleich mit der Mittelschwelle der westalpinen Geosyklinale der Jurazeit zu deuten gesucht hat[1]).

Gewiß arbeitet die moderne Morphologie, den Bahnen folgend, die ihr insbesondere Ferdinand von Richthofen gewiesen

---

1) E. Haug, Traité de Géologie, Teil I, S. 164. Paris, 1907.

hat, retrospektiv in dem Sinne, daß sie das örtliche Landschaftsbild aus dem Baue seines Untergrundes und damit aus der in diesem Baue sich ausdrückenden geologischen Vergangenheit zu erklären sucht; aber nicht das habe ich im Sinne, wenn ich von retrospektiver Behandlung von Problemen der heutigen Erdgestaltung spreche, denn in dem oben erwähnten Falle der Meeresräume und in anderen mir vorschwebenden Fällen ist der Untergrund unserer Betrachtung überhaupt nicht oder doch nur lückenhaft zugänglich; vielmehr handelt es sich um eine vergleichende Rückschau nach analogen Verhältnissen, die in anderen Räumen in der geologischen Vergangenheit obgewaltet haben.

Auf diesem retrospektiven Wege möchte ich mich mit den sogenannten Tiefseegräben, den bedeutendsten Depressionen unserer Erdrinde, und mit ihren Beziehungen zu den von ihnen begleiteten Gebirgs- und Inselzügen befassen. Wir halten Umschau nach analogen Gebilden der Vorzeit, die uns zugänglich sind, und ziehen aus den zu machenden Feststellungen die Nutzanwendung für die Erklärung der rezenten Tiefen. —

I. Bisherige Deutung der Tiefseegräben.

Schon in den ersten Zeiten der Tiefseeforschung hatte man erkannt, daß die größten Tiefen nicht, wie man zunächst wohl hätte annehmen können, inmitten der Ozeane liegen, sondern dicht neben den Kontinentalrändern und Inselzügen. Im allgemeinen bilden sie hier schmale, langgestreckte Rinnen, die als charakteristische Bodenform von weiterer Verbreitung zuerst von Supan[1]) im Jahre 1899 erkannt worden sind. Supan schlug die Bezeichnung „Gräben" vor, wollte damit aber zunächst nur die Form und nicht dis Entstehungsart zum Ausdruck bringen. Allerdings fügte er hinzu, daß diese submarinen „Gräben" wohl als „Verwerfungserscheinungen im großen Stile" zu deuten seien.

Die größte bisher gelotete Tiefe von 9788 m, die Planettiefe, liegt im Philippinengraben nordöstlich von Nord-Mindanao, die 9636 m betragende Nerotiefe im Mariannengraben südlich der Marianneninsel Guam, die mit 9427 m festgestellte Pinguintiefe im Kermandec-Graben südlich der Tongainseln. Die Tuscaroratiefe, die mit 8500 m lange Zeit die größte bekannte Meerestiefe gewesen ist, fällt in den Japan-Graben. Überhaupt sind Tiefseegräben bis-

1) Supan: Die Bodenformen des Weltmeeres. Petermanns geogr. Mitt. Bd. 45, 1899, S. 177 ff.

her namentlich aus der peripheren Zone des Pazifik und hier wieder ganz besonders aus der westlichen Peripherie bekannt geworden; so entfallen 16 der 22 Gräben, die Krümmel[1]) 1907 kannte, in den Pazifik. So kennen wir, um den schon genannten pazifischen Gräben noch einige andere hinzuzufügen, den Liukiugraben östlich der Liukiu-Inseln, den Boningraben östlich der Bonininseln, den Palaugraben östlich der Palauinseln, den Yapgraben östlich der Insel Yap, den Hebridengraben westlich der Neuen Hebriden und den Tongagraben, der östlich der Tongainseln die nördliche Fortsetzung des oben erwähnten Kermandecgrabens bildet. Über den Alëutengraben südlich der Alëuten kommen wir zur östlichen Peripherie des Pazifik, wo Tiefseegräben den Westrand Südamerikas begleiten und wo im Querschnitte von der Atakamatiefe (— 7636 m) zum Vulkan Llullaillaco (+ 6600 m) der größte bekannte Höhenabfall der Erde von 14236 m auftritt.

In der vorwiegend von Küsten atlantischer Struktur begleiteten Peripherie des Indischen uud Atlantischen Ozeans sind Tiefseegräben selten; nur dort, wo ausnahmsweise pazifischer Küstentypus auftritt, finden wir sie, so in der Peripherie des Indischen Ozeans im Bereiche des Malayischen Archipels und im Atlantischen Ozean im Bereiche der Antillen. Eine einzige große Tiefe macht hinsichtlich ihrer Lage zum aufragenden Lande, aber auch hinsichtlich ihrer gedrungenen Form, eine Ausnahme, nämlich die bis zu 7370 m gehende Romanche-Tiefe inmitten des Atlantik.

Was nun die Entstehung dieser Tiefsee„gräben“ anlangt, so herrscht im allgemeinen heute noch die Vorstellung, die Supan im Jahre 1899 mit großer Zurückhaltung geäußert hatte, die er dann aber selbst wieder, wie wir sehen werden, aufgab, nämlich diejenige, daß die morphologischen Gräben auch in tektonischer Hinsicht „Gräben“, d. h. Versenkungen an Bruchlinien, seien. So hatte auch Krümmel (l. c.), wenn er auch zunächst nur die zurückhaltend geäußerte Ansicht Supans mit gleicher Zurückhaltung wiederholte, hinzugefügt, daß die Gräben offenbar den „Zerrungsbrüchen“ entlang den Küsten von pazifischem Typus angehören, und dabei war er wohl durch die Vorstellungen v. Richthofens über Zerrungsbrüche und Zerrungsbögen beeinflußt. Ganz besonders hat Eduard Sueß die Entstehung der Tiefseegräben durch radialen Einbruch vertreten und diese Vorstellung zur weitgehendsten Anerkennung gebracht.

---

1) Krümmel, Handbuch der Ozeanographie, 1907.

. Schon im 2. Bande des „Antlitzes der Erde“, der 1888 erschien, — also zu einer Zeit, in der man zwar schon von vereinzelten besonders großen Tiefen in der Nähe der Festländer und Inselketten wußte, aber doch die Tiefseegräben als allgemeineres Phänomen noch nicht erkannt hatte, — gilt für Sueß „das Auftreten großer Tiefen knapp an dem Rande der Gebirgsbögen, z. B. an der Küste von Japan und an der südamerikanischen Westküste“, geradezu als Hinweis darauf, daß die Meeresbecken Senkungsgebiete seien“, „in weit größerem Maßstabe jene Senkungen wiederholend, welche wir im Innern der Festländer kennen gelernt haben“. Besonders ist dann im 2. Teile des 3. Bandes des „Antlitzes der Erde“ (1909) von den Tiefseegräben die Rede, namentlich auch hinsichtlich ihrer Beziehungen zu den Faltenzügen, denen sie vorgelagert sind. Sie sind nach Sueß bis auf eine oder zwei Ausnahmen „Vortiefen“, d. h. Tiefen vor der konvexen Außenseite der Gebirgsbögen. Dabei sind sie (S. 722) „keine Synklinalen im tektonischen Sinne“, vielmehr resultieren sie „aus den radialen Elementen der Kontraktion, d. h. aus Senkung infolge Verminderung des planetarischen Volumens“. Ihre Ränder sind sehr ungleich, denn der äußere Rand ist der bogenförmige Rand einer Senkung der Lithosphäre und der innere Rand ist der Außenrand des Faltengebirges, welches vom Lande her über diese Tiefe getreten ist (S. 582). Das Verhältnis von Vortiefe und Gebirgsbogen wird dabei, wie auch schon E. Tietze[1]) hervorgehoben hat, nicht ganz einheitlich aufgefaßt. Einerseits handelt es sich um ein „Absinken des Vorlandes unter die jungen Falten“ (S. 336); „fast möchte man meinen, das Vorland werde durch das vordringende Gebirge eingedrückt“ (S. 722). Anderseits sollen sich die Falten bogenförmig in und über die gesenkten Vortiefen hinwegschieben; „die tangentiale Kraft bringt ihren Überschuß in bogenförmigen Falten in und über die gesenkten Vortiefen“ (S. 670). Einerseits soll also die Vortiefe die primäre Erscheinung sein und sollen sich die Falten über die vorhandene Vortiefe hinweg und in sie hinein bewegen, anderseits erscheinen die Falten als das primäre Gebilde und das Absinken des Vorlandes als die Folgeerscheinung. Wir haben Sueß wohl dahin zu verstehen, daß teilweise schon gefaltete Ketten bei Fortdauer der Faltung in die sich senkenden Vortiefen hinuntergleiten, d. h. daß Faltung und Vortiefenbildung nebeneinander hergehen. Verweist doch Sueß (S. 722) auf Middlemiss, der schon

---

1) E. Tietze, Einige Seiten über Eduard Sueß. Jahrb. d. K. K. Geol. Reichsanst. Wien, 1916 Bd. LVI, S. 333 ff., bes. S. 475 ff.

1890 angenommen hatte, daß die Senkung im Vorlande des Himalaya und die Vorfaltung nach dieser hin gleichsam Hand in Hand gegangen seien.

Die abyssischen Tiefen sind also für Sueß zwar keine Gräben im gewöhnlichen Sinne, „denn ihre Seiten sind von verschiedenem Bau", und darum hat Sueß die vorher von ihm gebrauchte Bezeichnung „Vorgräben" durch „Vortiefen" ersetzt (S. 336). Im Grunde hält er sie aber, wenn auch die tangentiale Bewegung der Falten zu den Vortiefen hinzukommt, für das Ergebnis radialen Einbruches.

Der Sueß'schen Vorstellung vom Vorrücken der Inselbögen gegen die Vortiefen stimmt E. Horn[1]) zu, ohne sich im übrigen zur Frage der Entstehung der Vortiefen zu äußern. Sodann findet sich der Gedanke der allmählichen Vorwärtsbewegung von Zonen geantiklinaler Art gegen Zonen von geosynklinalem Charakter bei Argand[2]) und Brouwer[3]), wenn auch nicht in ausgesprochener Anwendung auf die abyssischen Tiefseegräben, so doch in Anwendung auf in mancherlei Hinsicht vergleichbare Verhältnisse im malayischen Inselmeere und auf die nach Argand diesen Verhältnissen wieder analogen Erscheinungen in den mesozoischen Geosynklinal- und Geantiklinalzonen des heutigen Alpengebietes.

Es sind aber auch bereits Meinungen geäußert worden, die von den von E. Sueß im „Antlitz der Erde" vertretenen Auffassungen wesentlich abweichen.

Supan hatte, wie wir sahen, im Jahre 1899 die Bezeichnung Gräben rein morphologisch verstehen wollen, wenn er es auch für wahrscheinlich hielt, daß diese morphologischen Gräben zugleich Gräben in tektonischer Hinsicht seien. Im Jahre 1907 beschäftigt er sich zum zweiten Male mit dieser Frage in einer kurzen Mitteilung über die Sundagräben[4]) und verweist auf einen neuen Typus, den „Doppelgraben", der durch die Messungen des deutschen Vermessungsschiffes Planet festgestellt worden war. Südlich von Java liegt zunächst der Sunda-Innengraben, dann folgt eine submarine Bodenschwelle, die die östliche Fortsetzung der Mentawei-Inseln bildet, und erst südlich dieser Schwelle findet sich der

1) E. Horn, Über die geologische Bedeutung der Tiefseegräben. Geolog. Rundschau 1914, Bd. 5, S. 422 ff., insbes. S. 429.

2) E. Argand, Sur l'arc des Alpes occidentales. Eclog. geol. Helv. XIV, S. 145 ff.

3) H. A. Brouwer, Über Gebirgsbildung und Vulkanismus in den Molukken. Geol. Rundschau 1917, Bd. 8, S. 197—209.

4) Supan, Die Sundagräben. Peterm. Geogr. Mitt. 1907, Bd. 53, S. 70/71.

Hauptgraben (Sunda-Außengraben) mit der größten bisher bekannten Tiefe des Indischen Ozeans (7000 m). Supan gewinnt bei diesem Doppelgraben die Vorstellung, „als hätten wir es mit Faltungserscheinungen zu tun. Die Gräben wären in diesem Sinne als Synklinalen aufzufassen“. 1908 macht A. Penck[1]) geltend, daß wir „nicht den leisesten Anhalt“ dafür haben, daß die Tiefseegräben „ebenso durch Einbruch entstanden sind, wie der große Rheintal-Graben zwischen Schwarzwald und Vogesen“. Er vergleicht sie „dem Wellentale einer großen Grundfalte“. Unter Hinweis auf Supan meint dann auch H. A. Brouwer (l. c. S. 201), daß die langgestreckten Tiefseegräben Ostasiens „wahrscheinlich mit Faltungserscheinungen in ursächlichem Verband stehen“, ohne auf diese Frage weiter einzugehen.

Nach E. Haug[2]) sind die abyssischen Gräben des Pazifik die letzten Reste der peripazifischen Geosynklinale, die nach ihm seit der mesozoischen Zeit bestanden hat; am Ende des Tertiärs sollen die peripazifischen Ketten im Zuge der Geosynklinale entstanden sein und die Geosynklinale bis auf ihre Randzone verdrängt haben. In der submarinen Mittelschwelle des Doppelgrabens der Sunda-Inseln glaubt Haug ein Analogon zu der von ihm früher erkannten Mittelschwelle der jurassischen Geosynklinale der heutigen Westalpen zu erblicken.

Ein Widerspruch gegen die Sueß'schen Vorstellungen ist ferner von P. Marshall[3]) erhoben worden, und zwar in Bezug auf den Kermandec-Tonga-Graben, der sich von Neu-Seeland nach Samoa östlich der „Neuseeland-Kette“ Dana's, d. h. der Hochzone Neu-Seeland—Kermandec-Inseln—Tonga-Inseln, hinzieht. Sueß hatte angenommen (Antlitz III, 2. S. 349/358), daß die Ruahine-Kette am Ostrande der Neuseeländischen Nordinsel nach Osten bezw. Nordosten, d. h. also in der Richtung auf die Tiefe, in junger Zeit gefaltet sei. Demgegenüber ist nach Marshall die Faltung hier vorkänozoisch, denn das Tertiär ist nach ihm in der Hauptmasse nicht gefaltet, wohl aber 4000 m über den Meeresspiegel gehoben.

---

1) Albr. Penck, Die Entstehung der Alpen. Ztschr. Ges. f. Erdk. 1908.

2) E. Haug, Traité de Géol. 1907, p. 165. Vgl. auch p. 350: „Diese Gräben sind Geosynklinalen vergleichbar, die am Rande der mesozoischen Geosynklinale bis zur heutigen Zeit persistiert haben“,

oder S. 528: „Von der alten Geosynklinale bestehen noch bestimmte Tiefen, die unzusammenhängende Depressionen bilden. Das sind die heutigen Gräben des Mittelmeeres, des Malayischen Archipels, des Sundameeres und der Randzone des Pazifik“.

3) P. Marshall, Ozeania. Handb. d. reg. Geol. VII, 2, 1911, p. 28.

Nach Marshall entspricht die Erhebung der Kermandec-Tonga-Schwelle über der mittleren Tiefe des Pazifik etwa der Tiefenlage des Kermandec-Tonga-Grabens unter der mittleren Tiefe; für ihn liegt hier der Fall einer einfachen Anti- und Synklinale vor.

II. Versuch einer Deutung der Tiefseegräben auf vergleichend-retrospektivem Wege.

a) Fossile und recente Vortiefen.

E. Sueß[1]) hat wohl als erster auf die Analogie zwischen den abyssischen Tiefen, die vor den jungen pazifischen Faltungszonen liegen, und den verschütteten Vortiefen anderer junger Gebirge sowie den, wie ich sagen möchte, „fossilen" Vortiefen älterer Gebirgszüge hingewiesen. Er hebt hervor, daß den Faltenzügen tertiären Alters vielfach langgezogene Täler und Tiefen vorzuliegen scheinen, und verweist insbesondere auf das Tal des Guadalquivir im Norden der Betischen Cordillere, auf den Persischen Meerbusen und die Niederung des Euphrat und Tigris vor den Zagrosketten und auf das südliche Vorland des Himalaya. In letzterem Falle hebt er den merkwürdigen Parallelismus der südlichen Grenze der Niederung, in welcher der Ganges fließt, mit dem Fuße des Gebirges besonders hervor (Antlitz III, 2, S. 705). Mächtige Alluvionen, so die des Euphrat und Tigris im Vorlande der Zagrosketten oder des Ganges und Brahmaputra im Vorlande des Himalaya, arbeiten zwar an der Verhüllung der Vortiefen, — aber denken wir uns diese Alluvionen fortgenommen, so erscheint die tiefe Senkung des Landes. Auch die im südlichen Vorlande des Himalaya festgestellten Lotablenkungen unterstützen nach Sueß (l. c. S. 707) die Vorstellung der Vortiefe. Schon Burrard[2]) hatte gemeint, daß die allgemeine Ablenkung des Lotes gegen Süden eher durch Massenabgänge unter der Ebene als durch solche unter den Bergen veranlaßt sei. „In der Tat erklären sich alle durch das Lot ermittelten Angaben durch die Annahme, daß die Vortiefe mit Sediment von geringerer Dichte angefüllt ist, unter welches sich von Süden her die alten Felsarten des Vorlandes hinneigen, bis sie unter die steile Stirn des Himalaya gelangen. Das ist dieselbe

---

1) Vgl. insbes. Antlitz der Erde, III, 2, S. 335.

2) S. G. Burrard, Intensity and Direction of Force of Gravity in India. Phil. Trans. vol. 205 A. S. 289—318.

Vermutung, zu welcher vor der Stirn der Alpen die Erbohrung des böhmischen Granits bei Wels geführt hat. Es dürfte auch nicht ausgeschlossen sein, daß einst Berechnungen über die unterirdische Gestalt der Vortiefe möglich werden . .“ (Sueß, l. c., S. 707).

Die mächtige Sedimentation der Tertiärzeit vor dem Außenrande vieler unserer Hochgebirge bezeugt uns, daß wir auch hier starke Senkung gehabt haben. So befand sich vor den Alpen eine Vortiefe zur Zeit der Molasse, deren Mächtigkeit nach Alb. Heim[1]) vor den Schweizer Alpen nahe dem Gebirgsrande 2500 m, ja an manchen Stellen wohl über 3000 m beträgt. Im Norden wurde die Vortiefe durch die Böhmische Masse begrenzt, deren Südrand ziemlich steil unter die Molassezone einschiebt; wurde doch bei Wels südlich Linz, d. h. 17 km südlich des böhmischen Grundgebirges, der moldanubische Cordieritgneis unter 1037 m Tertiär erbohrt[2]). Ähnlich ist die Sachlage hinsichtlich des Raumes zwischen den Karpathen und der Podolischen Platte, in dem A. von Alth[3]) einen großen „Einsturz“ vermutet hatte, „wodurch ein mehrere Meilen breiter Abgrund gebildet wurde, den später die tertiären Absätze ausfüllten“ (vgl. Sueß, Antlitz I, S. 242). Eine ähnliche Vortiefe schied in der Tertiärzeit die Betische Cordillere von der Spanischen Meseta.

Vortiefen befanden sich in oberkarbonischer Zeit am Nordrande des variszisch-armorikanischen Gebirges von Schlesien über Westfalen, Rheinland, Belgien und Südengland bis hin nach Irland. In ihnen entstanden die flözreichen Bildungen des Oberkarbons, deren Tausende von Metern betragende Mächtigkeit Zeugnis von dem Ausmaße der Senkung ablegt. Wie sich nach Marcel Bertrand und E. Sueß das variscische Gebirge Europas in den paläozoischen Faltenzügen des östlichen Nordamerikas fortsetzt, so finden wir auch dort die Vortiefe als lange flözreiche Zone wieder.

Die rein äußerliche Ähnlichkeit zwischen den jungen Tiefseegräben und den fossilen Vortiefen ist allerdings gering. Denn im ersteren Falle handelt es sich um abyssische Tiefen, während im letzteren Falle die Senkungen mehr oder weniger aufgefüllt waren. Aber ob eine Zone tektonischer Senkung abyssisch ist und bleibt

1) Albert Heim, Geologie der Schweiz, Bd. 1, S. 40. Leipzig 1916.

2) E. Sueß, Über die Zerlegung der gebirgsbildenden Kraft. Verh. d. Geol. Ges. Wien, VI. Band, S. 13 ff.

3) A. von Alth, Über die paläozoischen Gebilde Podoliens etc. Abh. K. K. geol. Reichsanst. Wien, Bd. VII, S. 4. 1874.

oder ob sie aufgefüllt wird, hat nichts mit den endogenen Vorgängen ihrer Entstehung zu tun, sondern liegt in den exogenen Vorgängen der Sedimentation begründet. Wir müssen die „tektonische“ und die „morphologische“ Tiefe auseinanderhalten, und es ergibt sich dann die jeweilige morphologische Tiefe aus der Differenz von Senkungsbetrag (tektonische Tiefe) und Sedimentationsmächtigkeit. So kann die tektonische Tiefe sehr bedeutend sein, trotzdem eine morphologische Tiefe nicht oder kaum vorhanden ist, da eine mächtige Sedimentation den Senkungsbetrag ausgeglichen und dabei in der Mehrzahl der Fälle mit der Senkung wohl einigermaßen Schritt gehalten hat. Kennen wir doch Erdräume genug, die um Tausende von Metern allmählich gesunken und deren Oberflächen dabei niemals oder höchstens vorübergehend, wenn einmal die Sedimentation die Senkung nicht zu kompensieren vermochte, unter das Meeresniveau gelangt sind, oder andere, in denen sich trotz der Senkung um Tausende von Metern stets Sedimente von Flachseecharakter gebildet haben. Wo dagegen die Sedimentzuführung gering ist, sind bei entsprechender Senkung die Vorbedingungen zum Ozean oder gar zu abyssischen Ozeanstiefen erfüllt. In diesem Sinne kommt vielleicht in Betracht, daß die abyssischen Tiefen durchweg vor Küsten mit nur schmalen tributären Zonen, meist vor schmalen Inselzügen, liegen. Das ist im allgemeinen schon im Wesen der pazifischen Küstenstruktur, an die unsere abyssischen Rinnen geknüpft sind, begründet. Doch gibt es ausnahmsweise auch pazifische Küstenzonen mit weitem tributären Hinterlande; ich verweise z. B. auf die westliche Küstenregion Nordamerikas, der nun aber auch die Vorgräben fehlen. Wären wir imstande, vor unseren heutigen Küstengebirgen die „tektonische“ Tiefe auch dort zu erkennen, wo über ihr die morphologische Tiefe verloren gegangen oder auch garnicht zur Entwicklung gekommen ist, so würden wir erst die zutreffende Vorstellung von der Verbreitung des ganzen Vortiefenphänomens erhalten und jedenfalls erkennen, daß unsere abyssischen Vortiefen nur ein reduzierter, unter günstigen Umständen erhalten gebliebener Teil des weit allgemeiner verbreiteten Phänomens der tektonischen Vorsenken sind.

Aber nicht allein, daß die tributären Landzonen nur schmal sind, — sie sind in der allergrößten Zahl der Fälle auch nur einseitig vorhanden, während ihr „Vorland“, um mit Sueß zu reden, unter dem Meeresspiegel liegt. Und wo in vereinzelten Ausnahmefällen im pazifischen Inselmeere, wie südwestlich des Hebridengrabens, das Land jenseits der Vortiefe bald wieder aufsteigt, geschieht das wiederum nur zu schmalen tributären Zonen.

Mit diesen Ausführungen soll gewiß nicht gesagt sein, daß die Sedimentation überhaupt nicht an der Auffüllung der jungen abyssischen Tiefen arbeitet. Nicht allein, daß hier selbstverständlich pelagische Sedimentation stattfindet, sondern es gelangt auch etwas terrigener Schlick in die Gräben hinab und auch subaquatische Rutschungen an den Hängen der Senken mögen eine Rolle spielen[1]); aber eine verhältnismäßig unbedeutende Zuführung sedimentären Materials im Vergleich zu anderen küstennahen Zonen mit weitem tributären Hinterlande scheint vorzuliegen. Zwar wird selbst eine verhältnismäßig geringe Sedimentation mit der Zeit auch tiefe Rinnen ausfüllen, und wenn solche trotzdem noch bestehen, so haben wir wohl Ursache, sie für jugendliche Gebilde oder wenigstens für Gebilde erneuter Senkung in jugendlicher Zeit zu halten. So erscheint uns also die Auffüllung der Vorsenken als ein sekundärer, rein exogener Vorgang, und so können wir solche Vorsenken, die unter den besonderen Verhältnissen der Sedimentzuführung den abyssischen Charakter nie besessen oder doch jedenfalls schnell verloren haben, mit E. Sueß sehr wohl als unseren Tiefseegräben analoge Gebilde betrachten und sie bei der Erörterung der Ursache der Senkungen und des Zusammenhanges zwischen Gebirgsbogen und Vorsenke vergleichend heranziehen.

In höherem Maße scheint zwar etwas anderes die Analogie zwischen den abyssischen Gräben der Jetztzeit und den tertiären Vortiefen der Alpen, der Karpathen und des Himalaya oder der jungpalaeozoischen Vortiefe des variscischen Bogens zu beeinträchtigen. Jenseits dieser fossilen Vortiefen lag nämlich ein festländisches Vorland von Art einer „Masse“ im tektonischen Sinne, d. h. eines mehr oder weniger „erstarrten“ Gebietes mit älterer Faltungsstruktur; bei den rezenten Vortiefen liegt dagegen das Vorland im allgemeinen tief unter dem Meeresspiegel; und wenn im „Vorlande“ des Hebridengrabens die Neukaledonischen Inseln oder wenn — im Sinne von Sueß (Antlitz III, 2, S. 525) — im Vorlande der westindischen Bartlett-Tiefe die Cayman-Inseln aufragen, so handelt es sich hier um Reste junger Gebirgszüge, nicht aber um „Massen“. E. Sueß[2]) glaubt dieser Schwierigkeit dadurch Herr werden zu können, daß er unter dem nördlichen Pazifik, und zwar auf Grund des „Vordringens“ der konvexen Ge-

---

1) Vgl. E. Horn, l. c. S. 431 ff.

2) Vgl. Sueß, Antlitz I, S. 591: „Wie gegen die indische Halbinsel dringen scharende Bögen von Nordwest, Nord und Nordost gegen den nordpazifischen

birgs- und Inselbögen auf diesen hin, ein „überflutetes Vorland“ annimmt, — aber das sind doch etwas hypothetische Vorstellungen, hinsichtlich deren ich Sueß aus verschiedenen Gründen nicht ganz folgen möchte. Und deshalb gehe ich auch nicht so weit, daß ich in den rezenten Tiefseerinnen und den fossilen Vortiefen von der Art derjenigen der Alpen, Karpathen und des variscischen Bogens genetisch in jeder Hinsicht analoge Gebilde erblicke, vielmehr sehe ich die Analogie zunächst nur darin, daß Zonen besonders tiefer Absenkung den Rand von Gebirgsbögen, — bezw. von Inselzügen, die als Gebirgsbögen zu deuten sind —, begleiten.

Die Analogie geht weiter zwischen den alten Vortiefen und denjenigen jungen abyssischen Tiefen, zu denen hin die Faltung des angrenzenden Gebirgsbogens gerichtet ist, die also wirklich Vorgräben sind; hier erstreckt sich die Analogie nicht nur auf das Auftreten tiefer Senken neben Gebirgsketten, sondern auch auf ihre Lage vor diesen Gebirgsketten.

b) Die Geschichte der fossilen Vortiefen.

Wir betrachten nunmehr die Entstehung der alten Vortiefen und beginnen mit derjenigen des variscischen Gebirges.

Bekanntlich ist das variscische Gebirge nicht in einem Gusse entstanden, sondern es sind mehrere Phasen der alten Faltung nachweisbar, die mit Verbreiterungen des Gebirgsbogens verbunden gewesen sind. So haben wir eine ältere Kernzone des variscischen Gebirges, im Bereiche deren die Diskordanz im Liegenden des Oberkarbons die ältere Faltung bezeugt, von der jüngeren Randzone, in der sich nur örtlich (Oberschlesien) und dann auch nur ganz verschwächt der ältere Faltungsvorgang bemerkbar gemacht hat, in die im übrigen aber erst nach Ablagerung der oberkarbonischen Schichten die Faltung gewandert ist, zu unter-

---

Ozean vor. Es besteht eine ganz außerordentliche tektonische Homologie zwischen dem indischen Tafellande und dem nördlichen Teile des pazifischen Ozeans“.

Ferner Antlitz III, S. 581:

„Die indische Halbinsel ist Vorland . . . und wir müssen annehmen, daß auch unter dem nordpazifischen Ozean ein Vorland liegt. Man darf sich in der Beurteilung dieser Grundzüge des Baues nicht durch die Hydrosphäre täuschen lassen. Könnten wir sie entfernen, so würde ein weites Land vor uns liegen, etwa 3600—4000 m unter dem heutigen Meeresspiegel, aber vor den gefalteten Bögen würden langgestreckte, allerdings nicht völlig regelmäßige Streifen des Landes liegen, die noch um 2000, 3000 und sogar 5000 m tiefer wären als das weite neu entblößte Land. Dieses entblößte Land ist das Vorland“.

Ferner Antlitz III, 2, S. 692: „Der Grund des (pazifischen) Ozeans verhält sich wie ein überflutetes Vorland“.

scheiden. Diese jüngere Randzone war in der Zeit zwischen den beiden Faltungsphasen ein Gebiet besonders starker Sedimentation und, wie wir daraus schließen müssen, besonders starker Senkung, — sie war eben jene Vortiefe vor dem konvexen Außenrande des älteren und inneren Teiles des variscischen Bogens, auf die schon Sueß hingewiesen hat und in der immer wieder der Randschutt der im Süden liegenden Ketten in vielfacher Unterbrechung mit Kohlenbildungen zur Ablagerung gekommen war. In Belgien, Rheinland und Westfalen ist nun in ihr das Oberkarbon lückenlos und konkordant mit dem Unterkarbon verbunden, und die obercarbonische Sedimentation bezeugt den Fortgang des Senkungsprozesses, der hier auch schon im Devon und Unterkarbon eingetreten war; aber während die Senkung früher ein weit nach Süden sich anschließendes Gebiet mitumfaßt hatte oder sogar dort hauptsächlich eingetreten war, ist sie nach Auffaltung dieses Südgebietes, d. h. im Oberkarbon, auf den Norden beschränkt oder in ihm doch ungleich stärker. Von einer „Graben"bildung im tektonischen Sinne oder von sonstigen orogenetischen Vorgängen, die die Vortiefenregion abwärts bewegt hätten, kann hier keine Rede sein, denn nirgends sind Diskordanzen zwischen dem Unterkarbon und Oberkarbon oder innerhalb des Oberkarbons nachgewiesen. Überhaupt ist keinerlei Beweis gegeben, daß diese alte Vortiefe anders entstanden wäre und anders sich fortgebildet hätte, wie auch sonst die Räume von geosynklinalem Charakter, d. h. als eine Mulde von großer Spannweite. Das bezeugt auch die ganze Art der oberkarbonischen Sedimentation. Die Auffaltung dieses Gebietes ist erst nach Ablagerung des dortigen Oberkarbons, das die Saarbrücker Stufe noch umfaßt, erfolgt. Damals erneuerte sich die Faltung im Süden, ergriff nun aber auch das Gebiet der nördlichen Vortiefe und war, wie auch die ältere Faltung, nordwärts, d. h. von dem alten Gebirgsbogen zur Vortiefe hin, gerichtet.

Wir haben also zu unterscheiden

1) die ältere Nordwärtsfaltung im Süden der späteren Vortiefe (zwischen Unterkarbon und unterem Oberkarbon),
2) die Vortiefenzeit, in der sich am Nordrande des älteren Gebirges die Einsenkung und die Auffüllung dieser Einsenkung vollzog (unteres und mittleres Oberkarbon),
3) die jüngere Nordwärtsfaltung, die auch die Vortiefe ergriff (nach dem mittleren Oberkarbon).

Auf die Vortiefenzeit kommt es hier besonders an; in ihr bestand also eine Vortiefe vor einem Gebirgsbogen, dessen Faltungen und Überschiebungen auf die Vortiefe gerichtet waren,

wie auch zu den heutigen Vortiefen die Falten und Überschiebungen gerichtet sind. Aber nirgends ist innerhalb dieser langen Zeit auch nur eine Andeutung von Faltungsvorgängen oder von irgendwelcher Massenbewegung in der Richtung auf die Vortiefe gegeben, vielmehr erfolgten ausschließlich Bewegungen von epirogenetischer Art und es ruhten die orogenetischen Kräfte.

Ganz analog liegen die Verhältnisse in der subalpinen Geosynklinale, der miocänen Vortiefe der Alpen, die schon Marcel Bertrand[1]) mit dem Kohlengürtel am Außenrande des variscisch-armorikanischen Bogens verglichen hat.

Wir haben eine ältere Faltung vor der Molassezeit. Auf sie geht die Struktur der südlich der Molasse liegenden Alpenzonen zu einem wesentlichen Teile zurück.

Wir haben eine lange Periode ausschließlich epirogenetischer[2]) Vorgänge, während deren das Molassegebiet sich im Zustande der Abwärtsbewegung befand und etwa entsprechend dieser Abwärtsbewegung eine Auffüllung erfuhr. Im allgemeinen ging die Ablagerung der Molasse kontinuierlich vor sich und nur örtlich traten Unterbrechungen und Erosionen mit nachfolgender Transgression ein. Dabei war die Sedimentation z. T. festländischer, z. T. mariner Art. Die Auffüllung erfolgte in der Hauptsache durch alpinen Randschutt, der von Süden her in die Vortiefe gelangte. Aber auch das nördlich liegende Gebiet lieferte Material und natürlich besonders für die nördliche Randzone des Molassegebietes. So kennen wir ja im Schweizer Jura Gerölle alpiner Herkunft nur in den Konglomeraten des Vindoboniens, in der sog. „bunten" Nagelfluh, während im übrigen die Gerölle der dortigen Molasse nördlicher Herkunft sind.

---

1) Marcel Bertrand, La Chaîne des Alpes et la formation du continent européen. Bull. Soc. Géol. de France, 3. Série, Bd. XV.

2) Ich fasse den Begriff epirogenetisch hier nicht auf, wie E. Haug, der alle tektonischen Bewegungen in den Geosynklinalzonen, wenigstens soweit sie posthum zur älteren Struktur gerichtet sind, als orogenetisch anspricht, während nach ihm die epirogenetischen Vorgänge an die Extrageosynklinalgebiete gebunden sind und ihr wesentliches Kennzeichen in der zur Struktur des Untergrundes transversalen Richtung liegt; vielmehr gebrauche ich, wie ich immer getan habe, ganz unbekümmert darum, ob es sich um posthume oder renegante Vorgänge und um solche in Geosynklinal- oder Extrageosynklinalgebieten handelt, den Begriff epirogenetisch, — m. E. in voller Übereinstimmung mit der Gilbert'schen Urbedeutung dieses Begriffes —, für den ganzen Komplex der „säkulären" Bewegungen, auch wenn sie, wie bei dem Gilbert'schen Urtypus der epirogenetischen Gebilde, dem Great Basin, abwärts gerichtet sind. Davon ein andermal.

Wir haben sodann Gebirgsbildungen nach der Molassezeit und zwar, wie wir nach Analogie der Verhältnisse in anderen Teilen des Alpenbogens wohl annehmen dürfen, vor Ablagerung der Pontischen Stufe, d. h. also im Ausgange des Miocäns.

Es scheint auch die Ansicht von E. Sueß zu sein, daß im Falle der fossilen Vortiefen die Faltung nach Entstehung der Vortiefen eingetreten und daß also das Vorrücken der Faltung auf die Vortiefe eine „spätere Veränderung“ ist. „Die flözreiche Zone am Nordrande der westlichen Altaiden von Schlesien bis Irland und jenseits des Ozeans bis über den Mississippi, die Flysch- und die Molassezone im Norden der Alpen, die Siwalikzone im Süden des Himalaya sind in der Hauptsache Einlagerungen in Vortiefen gewesen. Ihre spätere Veränderung hat darin bestanden, daß der dem Gebirge zunächst liegende Teil von diesem überschritten und eingefalten wurde, sodaß z. B. das belgische Flözrevier als eine zerdrückte und listrisch zerteilte Synklinale über die älteren Gesteine im Süden gelangt“ (Antlitz III, 2, S. 722).

In Bezug auf das variscische Gebirge wird wohl allgemein angenommen, daß zur Zeit der Entstehung der Steinkohlenflöze Westfalens u. s. w., also in der Vortiefenzeit zwischen der vorobercarbonen und der jüngeren Faltung, keine orogenetischen Bewegungen, sondern ausschließlich senkende und hebende Bewegungen von epirogenetischer Art, mögen sie am Ende auch auf einen weitspannigen Faltenwurf hinauskommen, eingetreten sind, und in Bezug auf die Verhältnisse des variscischen Gebirges ist damit auch das Unterbleiben von Faltung und Vorschub in der Richtung auf die Vortiefe in der langen Zeit des Bestehens und der Fortbildung der Vortiefe wohl die allgemeine Auffassung. Dagegen ist die Auffassung in Bezug auf die alpine Vortiefe der Miocänzeit nicht derartig einheitlich, vielmehr wird hier ein mehr oder weniger kontinuierliches Vorgleiten des Alpenbogens in der Richtung auf die Vortiefe gelegentlich behauptet, und zwar im Zusammenhange mit der allgemeineren Vorstellung einer angeblich „kontinuierlichen“ Gebirgsbildung. Der richtige Kern dessen liegt darin, daß die tektonischen Kräfte, wenn wir unter „tektonisch“ nicht nur die orogenen, sondern auch die epirogenen Vorgänge verstehen, kontinuierlich wirken; aber wir müssen die beiden Arten von Vorgängen auseinander halten, die einen, die zu Änderungen in der tektonischen Innenstruktur des Untergrundes (Faltungen, Überschiebungen, Verwerfungen) führen, die anderen, die in den langsamen Vorgängen der Hebung oder Senkung gewisser Zonen zum Ausdrucke kommen, — und tuen wir das, so finden wir, daß, wie

das variscische Gebirge, so auch die Alpen nur in verhältnismäßig wenigen und engumgrenzten Zeiten, von denen eine z. B. vor der Molassezeit, eine andere am Ausgange des Miocäns liegt, Faltungen und Deckenüberschiebungen erfahren haben, während für die Verknüpfung der dazwischen liegenden säculären Vorgänge (Sinken der Geosynklinalen, Hebung der Geantiklinalen) mit Faltungs- oder Überschiebungserscheinungen, wie auch mit Verwerfungsvorgängen, nirgends ein wirklicher Beweis erbracht zu sein scheint. Erst weise man einmal in der Randzone der alten Vortiefe gegen den Gebirgsbogen die Winkeldiskordanzen nach, die bei vielfacher Vorwärtsbewegung des Gebirges in der Zeit der Sedimentation der Molasse innerhalb der Molassesedimente hätten entstehen müssen. Aber nichts derartiges ist bekannt geworden[1]).

Was nun das Verhältnis von Bogenform des Gebirges und Vortiefe bei den alten Gebirgen anlangt, so ist nach obigen Ausführungen die Bogenform die primäre Erscheinung. Sie bedingte also die Bogenform der Vortiefe, nicht ist es umgekehrt, nicht kam die Bogenform des Gebirges durch das Vorgleiten der entstehenden Falten in die Vortiefe hinein zustande.

Zu folgenden Ergebnissen sind wir also hinsichtlich der fossilen Vortiefen gekommen:

1) Die fossilen Vortiefen sind keine Gräben im tektonischen Sinne, d. h. keine Einbrüche an Verwerfungen, sondern sie sind Zonen bruchloser und säkulärer (epirogenetischer) Absenkung am Außenrande praeexistierender Gebirgsbögen.

2) In der langen Zeit des Einsinkens der Vortiefen, das durch die Mächtigkeit und Art der Sedimentation bezeugt wird, ruhten die orogenetischen Vorgänge. So sind auch Vorwärtsbewegungen der Falten zu den Vortiefen in den langen Zeiten der Fortbildung der Vortiefen nirgends erweisbar.

3) Erst in einer jüngeren orogenetischen Phase, die eine Erneuerung und ein Vorwandern der Faltung brachte, erfolgte die Faltung in der Richtung auf die Vortiefe, in dieser selbst und auch über sie hinweg.

---

1) Und sollte wirklich innerhalb der Molassezeit eine einzelne orogenetische Bewegungsphase nachgewiesen werden, so würde das an der grundsätzlichen Auffassung über die Episodicität der Bewegung gegen die Vortiefen und überhaupt der Alpenfaltung nichts ändern. Nur würde sich die miocäne Vortiefenzeit in zwei Unterzeiten zerlegen, für deren jede aber das Unterbleiben von Vorwärtsbewegungen des Alpenbogens zuträfe.

4) Die Bogenform der Vortiefe ist die Folge der Bogenform des Gebirgsbogens, vor dem die Vortiefe entstand, nicht ist die Sachlage umgekehrt.

Wenn Sueß (Antlitz III, 2, S. 600/601) sagt, daß „Ausfüllung der Tiefe, allgemeine Abtragung und das Vortreten des Gebirges auf dem Festlande gemeinschaftlich an der Verhüllung der Vortiefen arbeiten“, so soll wohl das „gemeinschaftlich“ nicht im Sinne von „gleichzeitig“ verstanden werden, denn das „Vortreten des Gebirges“ ist ein jüngerer Akt gegenüber der „Ausfüllung der Tiefe“ und der „allgemeinen Abtragung“, der eintritt, wenn die Vortiefen bereits gänzlich oder bis auf geringe Reste verhüllt sind, und der dann mit der gänzlichen oder teilweisen Auffaltung des Vortiefenraumes verknüpft ist. Und wie überhaupt einseitig vorstrebende Falten über das, was vor ihnen liegt, hinweggehen, so schieben sich die Falten bei eintretender Orogenese über Teile der Vortiefen hinweg und verhüllen nunmehr einen Teil des alten Vortiefenraumes. Hält man sich einen Inselbogen und die vor ihm liegende tiefe Depression vor Augen, so liegt es vielleicht nahe, für den Fall einer wiedereinsetzenden Faltung ein Hingleiten oder gar Hinabgleiten der Falten zu dem Tiefenraume etwa im Sinne der Vorstellungen Reyer's anzunehmen; aber die uns jetzt beschäftigenden fossilen Vortiefen waren bei Eintritt der Faltung nicht mehr als tiefe Depressionen vorhanden, sondern bereits aufgefüllt, — ganz abgesehen davon, daß mit der erneuten Faltung ja auch der Vortiefenraum gefaltet und hochbewegt wurde. Somit kann ein besonderes Bestreben auf Überfaltung der Vortiefen nur in dem Sinne, wie überhaupt ein einseitiges Vorwärtsdrängen von Falten und Schubdecken erkennbar ist, zugegeben werden. Jedenfalls zeigt es sich aber nicht während der Fortbildung der Vortiefen, sondern bei den nachfolgenden Faltungsakten.

c) Nutzanwendung aus der Geschichte der fossilen Vortiefen auf die heutigen Vortiefen.

Aus obigen Ausführungen ergibt sich folgende Nutzanwendung auf unsere heutigen Vortiefen:

1) Die Vortiefen sind keine Gräben im tektonischen Sinne, d. h. keine Versenkungen an Bruchspalten, überhaupt keine orogenetischen Gebilde, sondern Zonen verhältnismäßig starker säkulärer (epirogenetischer) Absenkung, d. h. Zonen von geosynklinalem Charakter. Sie sind Zonen gesteigerter

säkulärer Senkung in der Peripherie der säkulär sinkenden ozeanischen Räume.

2) Zuerst ist die Gebirgskette mit ihren nach außen gerichteten Falten und Überschiebungen da und dann entsteht an ihrem Außenrande die Vortiefe oder es verstärkt sich dort eine vorhandene Senkung bis zu dem Maße, daß von einer Vortiefe gesprochen werden kann. Der Gebirgsbogen ist also die primäre, die Vortiefe die sekundäre Erscheinung; da die nach außen gerichtete Falten- und Schuppenstruktur der Gebirgskette vorhanden ist, ehe die Vortiefe entsteht, so kann sie auch nicht durch ein Vorwärtsgleiten der Gesteinsmassen zu der Vortiefe hin bedingt sein.

3) Überhaupt ruhen im Gebirgsbogen während der Entstehung und Fortbildung der Vortiefe die orogenetischen Vorgänge, wie Faltungen und Überschiebungen, und nur Bewegungen von epirogenetischer Art treten ein, die den Gebirgsbogen in seiner Gesamtheit heben oder senken oder schwach verbiegen. Die Vorstellung, daß Faltung und Vortiefenbildung Hand in Hand gingen, und daß insbesondere in junger Zeit noch Bewegungen des Gebirgsbogens in der Richtung auf die Vortiefe erfolgt seien, ist damit abzulehnen.

4) Soweit unsere jungen abyssischen Vortiefen bogenförmig gestaltet sind, — was keineswegs in allen Fällen zutrifft —, ist ihre Bogenform die Folge der Bogenform der Gebirgszüge, an derem Außenrande sie liegen, — nicht aber ist umgekehrt die Bogenform der Gebirgszüge bedingt durch die Bogenform der Vortiefen etwa in dem mehrfach angenommenen Sinne, daß die Gebirgszüge zu den Vortiefen hin und in diese hinein vorglitten.

Wir sahen oben, daß sich namentlich in der asiatischen Inselwelt bezüglich des Verhaltens der Inselbögen zu den Vortiefen andere Auffassungen entwickelt haben, und es ist nun noch nötig, zu den dort für ein junges Vorrücken der Inselketten vorgebrachten Gründen Stellung zu nehmen.

Auf die Analogieen zwischen den heutigen Inselbögen Ostasiens und dem Zustande im Alpengebiete in der mesozoischen Zeit ist mehrfach hingewiesen worden. Haug[1]) hatte durch seine grundlegenden Untersuchungen über die faziellen Verhältnisse des Jura im Gebiete zwischen Gap und Digne eine Zerteilung der al-

1) E. Haug, Les chaînes subalpines entre Gap et Digne. Bull. Serv. Carte géol. Fr. III, No. 21, 1891. Vgl. auch E. Haug, Les géosynclinaux de la chaîne alpine pendant les temps secondaires. C. R. Acad. Sc., 14 juin 1909.

pinen Geosynklinale in Spezialgeosynklinalen unter Aufwölbung einer Mittelschwelle erkannt, und diese Vorstellung der Spezialgeosynklinalen und Spezialantiklinalen des Alpengebietes ist mehrfach bestätigt und erweitert worden. Ein Abbild dieser alten Zustände im Alpenbogen findet sich nach Argand[1]) u. a. in den bogenförmigen Inselzügen des Molukkengebietes und den sie trennenden Meerestiefen. Auf Grund dieser Analogie beeinflussen die Vorstellungen von der allmählichen Vorwärtsbewegung der Geantiklinalen und der als solche aufgefaßten Inselbögen gegen die Tiefen in dem einen und anderen Gebiete einander sehr, wobei vice versa das, was in dem einen Gebiete erkannt sein soll, zur Stütze der Auffassung im anderen Gebiete verwertet wird. Hier wie dort werden aber m. E. die episodisch „faltenden" und die säkulär „hebenden" und „senkenden" Vorgänge nicht ausreichend auseinander gehalten. Wohl hat das Molengraaf[2]) in Bezug auf den südostasiatischen Inselarchipel getan, indem er dort tektonische Vorgänge von zweierlei Art und zwar

1) solche von appalachischem Typus im Sinne von G. K. Gilbert und
2) solche von Basin-Range-Typus (Block-faulted-Type) im Sinne von G. K. Gilbert

unterschied. Zur ersteren Art gehört die große jungmiozäne Gebirgsbildung, die zu Deckenüberschiebungen, wie Wanner[3]) und Molengraaf sie von Timor beschrieben haben, geführt hat; die Vorgänge der zweiten Art schaffen nach Molengraaf die Inselzüge und die sie trennenden Meeresbecken. Dabei sind nach Molengraaf das Aufsteigen der Inselzüge und das Einsinken der Meeresbecken gleichzeitige und genetisch zusammenhängende Erscheinungen, hervorgerufen durch tangentialen Druck, — und zwar handelt es sich nach ihm um aufsteigende und sinkende große Blöcke, deren Bewegungen heute noch fortgehen.

In der Unterscheidung von zweierlei Art tektonischer Vorgänge stimme ich Molengraaf zu, nicht aber in der Deutung

---

1) l. c.

2) G. A. F. Molengraaf, On recent crustal movements in the island of Timor and their bearing on the geological history of the East-Indian Archipelago. Koninkl. Akad. van Wetenschappen te Amsterdam. 29. Juni 1912.

Derselbe, Folded Mountain Chaines, overthrust sheets and block-faulted mountains in the East-Indian Archipelago. Compte Rendu XII. Congrès Géolog. Intern. Toronto, 1913, S. 689 ff.

3) J. Wanner, Geologie von Westtimor. Geolog. Rundschau IV, 1913, S. 136 ff.

derjenigen der zweiten Art. Das Wesen des Block-Typus liegt in den Verschiebungen großer Blöcke entlang Verwerfungen und m. E. ferner, wie unter günstigen Verhältnissen in manchen Blockgebirgsgebieten nachweisbar ist, in der Kurzfristigkeit des Verwerfungsvorganges, — mag dieser kurzfristige Vorgang sich auch öfter wiederholen; der Block-Typus ist in diesem Sinne das Ergebnis einer Orogenese. Für die hebenden und senkenden Vorgänge, die Molengraaf im Malayischen Archipel vorschweben, ist aber die Verknüpfung mit Verwerfungen, d. h. die Verschiebung nach Art großer Blöcke, nicht erwiesen, vielmehr kommen wir durchaus mit der Annahme von Verbiegungen nach Art weitspannender Sättel und Mulden aus, und diese Verbiegungen bleiben trotz der Höhenunterschiede zwischen den Inselketten und den Meeresräumen, wie sich bei maßstäblicher Darstellung ergibt, ziemlich flach. M. E. haben wir hier 1.) kurzfristige orogenetische Vorgänge und zwar mindestens zwei Phasen derselben, nämlich a) solche jungmiocänen Alters, die zu starken Faltungen und, wie auf Timor, sogar zu Deckenüberschiebungen geführt haben, und b) solche spätpliocänen, vielleicht auch frühpleistocänen Alters, die das zwischen a) und b) abgelagerte Pliocän von Timor örtlich und insbesondere in seiner Randzone aufgerichtet haben. Im übrigen haben wir 2.) epirogenetische Vorgänge von langer Dauer, und zwar ältere, die zwischen den tektonischen Phasen a und b in der Fortbildung der das Pliocän von Mitteltimor umschließenden Senke zum Ausdruck kommen, und jüngere seit der orogenetischen Phase b), die auch heute noch andauern und deren Erkennung uns insbesondere die hochliegenden Korallenriffe, die auf Timor bis zu 1200 m ü. M. ansteigen, ermöglichen[1]).

---

1) Die Inselketten im malayischen oder pazifischen oder westindischen Inselmeere und damit auch die Inselketten neben den Vortiefen sind von Haus aus orogenetische Gebilde; aber sie werden, wie in der Erdgeschichte in so vielen ähnlich liegenden Fällen verfolgbar, ja sogar die Regel ist, wenn die Orogenese zonenweise versteifend gewirkt hat (vgl. Kap. III), zu Zonen von epirogenetischer Eigenbewegung — mag diese Eigenbewegung auch nur eine „Spezialundation" innerhalb eines größeren Raumes mit abwärtiger oder aufwärtiger „Allgemeinundation" sein. —

Durch positive Spezialundation bei negativer Allgemeinundation scheint mir überhaupt das Auftreten einzelner Hebungszonen im Bereiche sinkender Becken und auch der Wechsel zwischen Aufsteigen und Einsinken solcher Hebungszonen erklärbar zu sein. In seiner Gesamtheit ist z. B. das Gebiet der Molukken während der jüngsten Erdgeschichte ein sinkendes Becken; ragen hier doch nur noch einzelne Teile von Gebirgsketten als Reste einer hier vorhanden gewesenen ausgedehnteren festländischen Gebirgs-

Das ist nicht die „Fortdauer“ der faltenden Bewegungen der miocänen Zeit, sondern ein Vorgang ganz anderer Art, und vor allen Dingen ist es nicht, wie E. Horn (l. c.) meinte, die „unverminderte“ Fortdauer; was bedeutet in diesem Sinne eine in der langen Zeit seit etwa Ende des Pliocäns bis heute sich vollziehende Aufwärtsbewegung selbst um über 1000 m gegenüber den gewaltigen tektonischen Leistungen bei den kurzfristigen Faltungen und Überschiebungen, wie sie z. B. in der Struktur der Insel Timor zum Ausdruck kommen!

Hinsichtlich der von ihm behaupteten jugendlichen Vorwärtsbewegung der Gebirgsbögen in der Richtung auf die Vortiefen verweist E. Horn (l. c.) auf den Bau der Riu-Kiu-Inseln, die nach der Darstellung F. v. Richthofens[1]) Aufrichtung und Faltung gegen Südosten, d. h. in der Richtung auf die Vortiefe,

---

welt auf. Aber eine starke Spezialundation herrscht hier und treibt strichweise die Landmassen trotz der generellen abwärtigen Undation in die Höhe, wie die hochliegenden jugendlichen Korallenriffe beweisen. Diese Spezialundation ist aber nicht nur im Streichen, sondern auch quer dazu orientiert. Treten nun Veränderungen in der relativen Stärke von Allgemein- und Spezialundation ein, so kann es in denselben Zonen bald zum Aufsteigen (nämlich bei Überwiegen der positiven Spezialundation), bald zum Sinken (nämlich bei Überwiegen der negativen Allgemeinundation) kommen. Besonders tiefe Senkungen müssen aber dort auftreten, wo zu der negativen Allgemeinundation noch eine negative Spezialundation hinzukommt, und die allergrößten Senkungsbeträge müssen endlich, wie ich unter Bezugnahme auf die Ausführungen in Abschnitt III vorweg bemerke, dort erzielt werden, wo zu alledem sich noch eine besondere „Nahwirkung“ der aufsteigenden Inselketten geltend macht.

Als Teile eines im allgemeinen sinkenden Raumes sind die Inselbögen also ertrinkende Gebirge, die durch eigenen Auftrieb gegen das Ertrinken mit geringerem oder größerem, strichweise sogar mit starkem und andauerndem Erfolge ankämpfen.

Zu der Vorstellung dieser Wechselwirkung von Spezial- und Allgemeinundation bin ich auch durch die Verhältnisse im germanischen Triasbecken geführt worden. Immer tiefer sinkt das germanische Becken seit der Dyaszeit ein, wie die immer fortgehende Sedimentation in ihm bezeugt; das ist negative Spezialundation gegenüber den umrahmenden Schwellen, eine Spezialundation innerhalb des mitteleuropäischen Untergrundes variscischer Struktur. Aber nebenher gehen die Bewegungen, die das Becken samt seiner variscischen Umrahmung erfährt und die es mit dieser bald senken, bald heben, sodaß bald das Meer ingredieren kann, bald aber, — und zwar unter Fortgang der Spezialsenkung, — festländische Verhältnisse sich einstellen. Die Spezialundation bedingt also im germanischen Becken den Fortgang der Senkung, die Allgemeinundation die festländische oder marine Facies.

1) F. v. Richthofen, Geomorphol. Stud. aus Ostasien III. Sitz.-Ber. d. Kgl. Preuß. Akad. d. Wiss. 1902.

erkennen lassen. Ebenso sind in der Struktur der Neuen Hebriden die Äußerungen der gebirgsbildenden Kräfte gegen die Vortiefe, d. h. in diesem Falle gegen Westen, gerichtet; besonders zeigt sich hier das Tertiär der Außenzone stark gestört und gegen Westen steil aufgerichtet, gefaltet und überschoben. E. Horn überträgt nun die tangentiale Vorwärtsbewegung, die für tektonische Vorgänge älteren Datums, nämlich für die jungtertiäre Faltung, zutrifft, ohne weiteres auf ganz andersartige, nämlich vertikal sich äußernde Vorgänge der jungen geologischen Vergangenheit und Jetztzeit. Eine Stütze für seine Ansicht scheint er besonders darin erblicken zu wollen, daß in den Neuen Hebriden, die sich in der Vorwärtsbewegung gegen die westlich liegende Vortiefe befinden sollen, die jungen Korallenbildungen nach Westen ansteigen. Das heißt doch nur, daß die Aufwärtsbewegung im Westen stärker als im Osten ist, sodaß ursprünglich horizontale Riffe schräg gestellt wurden, nicht aber, daß hier eine Westbewegung des ganzen aufsteigenden Inselkörpers vorliegt. Ich erblicke hier nur den so häufigen Fall einer gewissen Verbiegung einer aufsteigenden Landschwelle [1]).

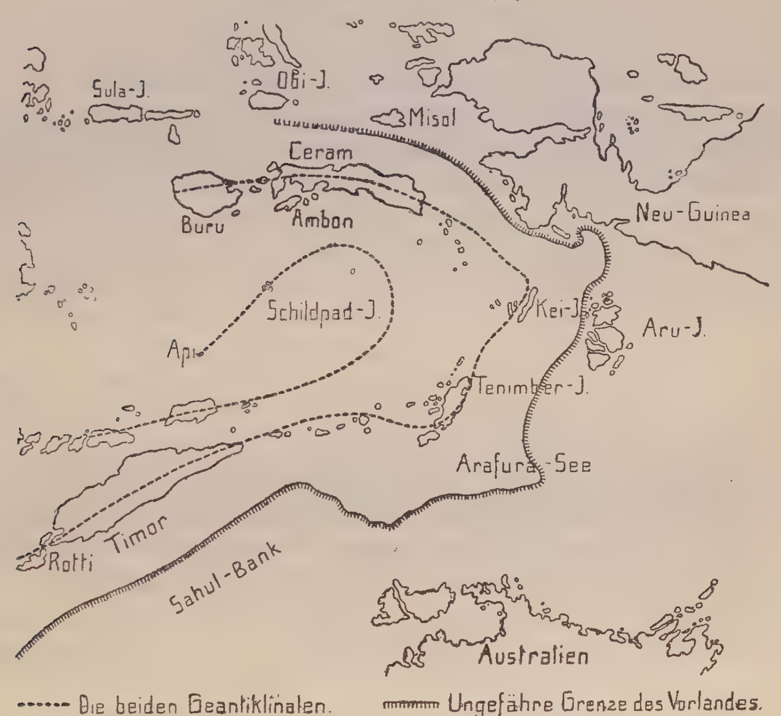
Ähnliche Gedankengänge wie bei Horn finden wir bei H. A. Brouwer (l. c.), und wenn bei Brouwer auch nicht ausdrücklich von der Vorwärtsbewegung von Inselbögen gegen abyssische Tiefseegräben, sondern ganz allgemein von ihrer Vorwärtsbewegung gegen angrenzende Meereszonen, die Rede ist, so möchte ich doch auf seine Ausführungen wegen ihrer unmittelbaren Konsequenz für die Verhältnisse der Tiefseegräben eingehen.

---

1) Auch Horn zieht (S. 436) vergleichend die Alpen heran, die nach ihm schon jetzt deutlich erkennen lassen, „daß sie nicht durch eine schnelle Auffaltung in kurzer Zeit entstanden sind, sondern daß sie im Laufe eines langen Entwicklungsprozesses, welcher mehrere geologische Perioden umfaßt hat, das Stadium der Inselbögen mit Vortiefe auf der Außenseite und Vulkanreihe auf der Innenseite durchgemacht haben“. Er stellt also „schnelle Auffaltung in kurzer Zeit“ und „langen Entwicklungsprozeß“, der auch das Stadium der Inselbögen umfaßt hat, einander gegenüber. Aber beides ist durchaus nicht unvereinbar. Die Alpen haben einen „langen Entwicklungsprozeß“ und dabei auch das „Stadium der Inselbögen mit Vortiefe“ durchgemacht, aber im Laufe dieses Entwicklungsprozesses ist es episodisch zu „schnellen Auffaltungen in kurzer Zeit“ gekommen, so an der Wende von Kreide und Tertiär, am Ende des Miocäns und nach der Molassezeit. Das ist aber (vgl. l. c. S. 447) keine „Kontinuiät“ der Gebirgsbildung, sondern das ist, da die epirogenetischen Vorgänge nicht unter den Begriff der Gebirgsbildung (Orogonese) fallen können, eine durchaus episodische Gebirgsbildung als Unterbrechung langzeitiger Vorgänge von rein epirogenetischer Art.

Brouwer behandelt insbesondere das in beistehendem Textbilde (Kopie nach Fig. 1 der Brouwer'schen Arbeit)[1]) dargestellte

Inselbögen und Vorland im Gebiete der Molukken.
Nach H. A. Brouwer.



Gebiet, das einerseits östlich von Java und Celebes und anderseits westlich von Australien und Neuguinea liegt und in dem tiefe Meeresbecken und gehobene Landmassen stark wechseln. Die gehobenen Landmassen gehören hier im wesentlichen zu zwei einigermaßen konzentrisch verlaufenden und durch tiefe Meeresbecken getrennten Inselbögen, dem äußeren Timor-Ceram-Bogen und dem inneren Bogen der jungen Vulkaninseln, der Endigung des Bogens Java-Soembawa-Flores-Wetter. Diese Inselbögen sind Faltungsbögen und die Faltung ist, soweit bekannt, nach außen gerichtet. Das Vorland der Falten bildet der in einzelnen Teilen das Meer über-

1) Nur die holländische Schreibart der Ortsbezeichnnngen ist durch die deutsche ersetzt und die auf das Vorland gerichteten Pfeile entlang den Inselzügen sind fortgelassen worden.

ragende, in anderen nur flach überflutete Sockel, der Australien einschließlich der Sahul-Bank, den östlichen Teil der Arafura-See, die Aru-Inseln, Neu-Guinea und einige Inseln westlich davon umfaßt. Dieses Vorland zeigt Einbiegungen und ihnen paßt sich der äußere Bogen weit mehr an, als der innere; in dieser Hinsicht wird von Brouwer besonders auf die Umbiegung des äußeren Bogens bei den Kei-Inseln, die ganz entsprechend der Kontur des Vorlandes erfolgt, hingewiesen. Die Kei-Inseln, die sich unmittelbar an den äußeren Bogen anlehnen, sind von der miocänen Faltung kaum noch betroffen, wohl aber äußern sich hier die „jüngsten gebirgsbildenden Vorgänge“ Brouwer's, d. h. die jungen Hebungen, besonders stark. Daraufhin spricht Brouwer von einem „Vorwandern der Faltung“ seit dem Miocän in der Richtung auf das Vorland und vergleicht sogar dieses angebliche Vorwandern, das gerade dort eingetreten sein soll, „wo im Vorlande eine Vertiefung vorkommt“, mit dem Vordringen der penninischen Überschiebungsdecken der Alpen in die Vertiefungen zwischen den autochthonen hercynischen Massiven, wie zwischen Montblanc- und Aarmassiv.

Ich kann zur Kritik dieser Brouwer'schen Auffassung nur das gleiche sagen, wie oben zur Kritik der von E. Horn angenommenen Vorwärtsbewegung der Inselzüge gegen die Vortiefen. Der sachliche Befund ist m. E. der, daß ein Gebirgsbogen, — in unserem Falle ein Inselbogen —, von älterem und dem Vorlande zugewandtem Faltenbau sich im Zustande der Aufwärtsbewegung befindet und daß an dieser Aufwärtsbewegung auch ein dem Bogen sich unmittelbar anlehnendes Erdstück teilnimmt, das von der älteren Faltung nur schwach betroffen war. Dieser Befund ergibt also eine ältere Faltung in der Richtung auf das Vorland und nachfolgende hebende Bewegungen. Aber den weiteren Schluß, daß auch diese jungen Vorgänge mit Vorwärtsbewegungen in der Richtung auf das Vorland verknüpft seien, sogar mit Vorwärtsbewegungen von Art alpiner Deckenschübe, lehne ich als unbewiesen ab. Daß die stärkere Anpassung des äußeren der beiden Bögen an das Vorland z. T. erst jungen Datums und z. T. erst zugleich mit der jugendlichen Aufwärtsbewegung erfolgt sei, erscheint mir unbewiesen und ist recht unwahrscheinlich gegenüber der Auffassung, daß der zur heutigen Inselreihe gewordene Gebirgszug schon von vornherein die ihm heute eigene Anpassung an seinen „Rahmen“, wie sie in der Umbiegung bei den Kei-Inseln oder auch in dem südlichen Vorspringen bei den Tenimber-Inseln zum Ausdruck kommt, erfahren hat. Von einem Vor-

wandern der „Faltung“ möchte ich hier aber gewiß nicht sprechen, denn die jungen Bewegungen, die sich auf den Kei-Inseln andeuten, sind keine „Faltungen“, sondern gehören in die Kategorie der säkulären Vorgänge, die nicht in Veränderungen der Struktur, sondern nur in vertikalen Bewegungen zum Ausdruck kommen; und daß diese jungen Bewegungen nun einmal besonders stark gegenüber einem Einsprunge des Vorlandes auftreten, halte ich umso weniger für einen Hinweis auf eine Verlegung der gebirgsbildenden Vorgänge vom Faltenzuge in der Richtung auf das Vorland, als im Faltenzuge selbst, so auf Timor, noch weit kräftigere junge Aufwärtsbewegungen eingetreten sind[1]).

d) Nutzanwendung aus der Geschichte der fossilen Vortiefen auf die Tiefseegräben im allgemeinen.

Die vorstehenden Vergleiche bezogen sich auf die fossilen Vorgräben von Art der subalpiden oder subkarpathischen oder subvariscischen einerseits und solche rezente Tiefseerinnen andererseits, die gleich diesen fossilen Vortiefen am Außenrande der

1) Während der Drucklegung erhalte ich eine inhaltreiche und in vielfacher Hinsicht sehr beachtenswerte Schrift von P. Arbenz über „Probleme der Sedimentation und ihre Beziehungen zur Gebirgsbildung in den Alpen“ (Heim-Festschrift. Vierteljahrsschrift der Naturf. Gesellsch. Zürich, Bd. LXIV, S. 246 ff.). In ihr vertritt, wie ich zunächst hervorhebe, auch Arbenz die Auffassung, daß Inselkränze und Tiefseegräben mit Geantiklinalen und Geosynklinalen zu vergleichen sind und zusammengehören „wie Gewölbe und Mulde“. Auch darin, daß beide nicht starr sind, sondern sich in Umbildung und Bewegung befinden, sind Arbenz und ich einer Meinung. Allerdings sagt er dann weiter, daß der Inselbogen gegen die Vortiefe vorrücke und sie zu verdrängen suche, und insbesondere hält er im Falle der Inselkränze des Molukkenmeeres die jugendliche und fortschreitende Bewegung gegen die Kontinentalränder und Schelfe von Australien und Neu-Guinea auf Grund der Ausführungen Brouwer's für feststehend. Das Analogon dieser Vorgänge erblickt er in den Westalpen; aber das „konkrete Beispiel“, das er aus der tertiären Geschichte der Westalpen für die Verdrängung alpiner Vortiefen durch angeblich fortschreitende Gebirgsbildung gibt, scheint mir eher im Sinne einer periodisch wiederkehrenden, als einer fortschreitenden Gebirgsbildung zu sprechen, Eine „Kontinuität der Deformation“ liegt hier selbstverständlich vor; aber die Deformationen durch Schub, — und nur auf diese, nicht aber auch auf die epirogenetischen Hebungen und Senkungen kommt es ja hier an —, sind in dem von Arbenz gewählten Beispiele, wie ich auch wieder aus Arbenz' eigenen Schilderungen glaube herauslesen zu müssen, wirklich erweisbar nur als episodische Ereignisse, die besonders im Ausgange der Kreidezeit, im Ausgange des Oligocäns und nach der Molassezeit eingetreten sind.

Gebirgs- oder Inselbögen liegen, und zunächst nur diese letzteren habe ich, wie oben nachdrücklich gesagt wurde, auf retrospektiv-vergleichender Weise erklären wollen. —

Aber wir haben auch junge abyssische Rinnen, die nicht Vortiefen der von ihnen begleitenden Kontinental- beziehungsweise Gebirgsränder sind, und ich verweise in diesem Sinne auf die südamerikanischen Rinnen, die sich auf der der Faltungsrichtung abgewandten, also auf der sogenannten Innenseite eines Gebirges, nämlich auf der Westseite der nach Osten gefalteten südamerikanischen Kordilleren, finden und in diesem Sinne eher als Rücktiefen bezeichnet werden könnten. Ich verweise ferner auf die oben schon erwähnten Ausführungen von P. Marshall, die eine junge Bewegung der Ruahani-Ketten Neuseelands in der Richtung auf den Kermandes-Tonga-Graben in Abrede stellen und damit auch die Eigenschaften dieses Grabens als einer Vortiefe in Zweifel ziehen. Auch in anderen Fällen fehlt noch der Beweis, daß die an die Vortiefen angrenzenden Gebirgszüge in der Richtung auf die Vortiefen gefaltet sind. In manchem solcher Fälle wird spätere Forschung klärend wirken können; in anderen, in denen nicht der gefaltete Sockel der Inselbögen, sondern nur die ihm aufgesetzten Vulkane und Korallenbauten das Meer überragen, wird aber der unmittelbare Beweis für eine Faltung in der Richtung auf die Vortiefen nie zu erbringen sein. —

Die ganze Erscheinung der abyssischen Tiefen an der Westseite des südamerikanischen Kontinents ist nun so übereinstimmend mit der Erscheinung solcher jungen abyssischen Tiefen, zu denen die Faltung hingewandt ist, daß man unmöglich an eine ganz andersartige Entstehung glauben kann, sondern daß man auch für sie die aus dem Werdegange der fossilen Vortiefen abgeleitete Bildung als Geosynklinalen annehmen muß. Und daß die übereinstimmende Erscheinung auftritt, mag die Faltung zu ihr hin oder von ihr weg gerichtet sein, beweist wieder, daß nicht die Faltungsrichtung in der benachbarten Gebirgskette, sondern höchstens die Existenz an sich und die Nachbarschaft dieser Gebirgskette für die Ausbildung der abyssischen Tiefen wesentlich sein kann, und das unterstützt die oben von den fossilen Vortiefen hergeleiteten Schlüsse gegen ein Handinhandgehen von Senkung der Tiefen und Vorfaltung der Gebirge. Alle die Ausführungen von E. Sueß in dem Sinne, daß die Ränder einer Vortiefe ungleichartig seien, nämlich der äußere ein Bruchrand, der innere der Überschiebungsrand des angrenzenden Faltengebirges, sind ja von vornherein ungültig für diejenigen

Tiefen, zu denen die Überschiebung oder Überfaltung nicht hingerichtet, sondern von denen sie weggerichtet ist; und es ist überraschend, daß E. Sueß die in diesem Sinne nur teilweise Gültigkeit seiner Ausführungen über die Tiefseegräben nicht hervorgehoben hat, trotzdem er auf die abweichende Lage der südamerikanischen Tiefen ausdrücklich hinwies. Hier bringt doch ganz gewiß nicht die tangentiale Kraft „ihren Überschuß in bogenförmigen Falten in und über die gesenkte Vortiefe" (siehe oben), sondern hier bringt sie ihn höchstens von der Tiefe weg. Ist nun aber der Innenrand der die Westküste Südamerikas begleitenden Tiefen kein Überschiebungsrand, so fällt es bei der Überstimmung des ganzen Phänomens schwer, für die Gestaltung des Innenrandes anderer abyssischer Tiefen das Überschiebungsphänomen als grundsätzlich bedeutungsvoll anzuerkennen —, und das deckt sich wieder mit den aus der Geschichte der fossilen Vortiefen abgeleiteten Schlüssen. Beide Ränder der Tiefseerinnen sind Senkungsränder, jedoch nicht die Ränder grabenartiger Versenkungen, d. h. Bruchränder, sondern geosynklinale Senkungsränder, die man höchstens als Flexurränder bezeichnen dürfte. Sie sind Erzeugnisse der säkulär wirkenden Undation, nicht aber orogenetische Erscheinungen irgendwelcher Art. —

Der Vergleich mit den fossilen Vortiefen zeigt also, daß die Vorstellungen, die Supan in Bezug auf die Sunda-Gräben, P. Marshall in Bezug auf den Kermandec-Tonga-Graben und wohl auch H. A. Brouwer in Bezug auf die ostasiatischen Tiefseegräben vorschwebten, im wesentlichen zutreffen dürften. Alle drei sprechen zwar ganz allgemein von „Faltungserscheinungen", haben aber wohl den weitspannigen säkulären (undatorischen) Wellenwurf im Auge, für den ich die Bezeichnung „Faltung", um Verwechslungen mit den episodischen eigentlichen Faltungen zu verhüten, vermieden sehen möchte. Die retrospektive Betrachtung führt uns aber vor allen Dingen zu der Definition zurück, die A. Penck schon vor elf Jahren, wenn auch ohne nähere Begründung, gegeben hat, indem er die Tiefseegräben als „Wellentäler einer großen Grundfalte" bezeichnete.

Der oben erläuterten Auffassung E. Haug's stimme ich insoweit zu, wie sie in den Tiefseegräben Gebilde von geosynklinaler Art sieht. Nicht aber kann ich die Auffassung der circumpacifischen Tiefseegräben als Restgeosynklinalen teilen. Ihr liegt die Vorstellung zu Grunde, daß alle Geosynklinalen verhältnismäßig schmale Senkungszonen zwischen zwei Kontinentalmassen seien und daß insbesondere eine circumpacifische Geosynklinale

nach innen von einem hypothetischen Festlande, das heute unter dem Pacifik versunken liegen soll, umgrenzt gewesen sei. Diese circumpacifische Geosynklinale, wie Haug sie sich in Konsequenz seiner allgemeinen Vorstellungen denkt, glaube ich aber mit vielen anderen Forschern ablehnen zu müssen, vielmehr betrachte ich den Pazifik ziemlich in seiner Gesamtheit als einen Ocean von hohem geologischen Alter. Für mich sind damit die circumpacifischen Tiefseegräben nicht der Rest einer im übrigen verschwundenen Geosynklinale, sondern sie bilden innerhalb eines uralten und in seiner allerweitesten Ausdehnung noch fortbestehenden Geosynklinalbeckens, nämlich des Pacifik, Zonen von zur Zeit besonders starker Vertiefung.

Das Zukunftsbild des heutigen abyssischen Raumes, — das lehrt uns das Schicksal der fossilen Vortiefen —, ist ein Gebirgszug in Angliederung an den schon bestehenden Gebirgsbogen. Aber deshalb möchte ich den Tiefseegraben nicht, wie mehrfach geschehen ist, als ein „werdendes“ Gebirge bezeichnen. Denn einstweilen ereignet sich hier das gerade Gegenteil einer Gebirgsbildung, nämlich keine Aufwärtsbewegung unter Veränderung der inneren Struktur (Faltung, Überschiebung etc.), sondern säculäre Abwärtsbewegung ohne Faltungen und Überschiebungen und sonstige „gebirgsbildende“ Erscheinungen. Es handelt sich hier nur um die Vorbereitung einer künftigen Gebirgsbildung, indem die Evolution eine für die spätere Erdrevolution prädestinierte Stätte schafft. Bei der Bezeichnung der Vortiefen oder der Inselbögen samt Vortiefen als „werdende Gebirge“ waltet die Vorstellung der bereits eintretenden horizontalen Vorbewegung der Inselbögen gegen die Vortiefen vor; aber das ist es gerade, was m. E. nicht zutrifft.

Was nun die Terminologie der tiefsten Rinnen unserer Weltmeere anlangt, so mag man die alte Bezeichnung „Tiefseegräben“ beibehalten, nur muß man sie rein beschreibend und nicht genetisch verstehen, wie Supan 1899 eigentlich auch gewollt hat (s. oben). Allerdings enthält dieser Ausdruck die erhebliche Gefahr seiner Verknüpfung mit der genetischen Vorstellung im Sinne eines Sueß'schen Grabens, denn die Bezeichnung einer rinnenartigen Depression als Graben ist sonst in der Morphologie nicht üblich; und auch von Supan wäre die Bezeichnung Tiefseegraben wohl kaum gebraucht worden, wenn er es nicht für wahrscheinlich gehalten hätte, „daß die submarinen Gräben Verwerfungserscheinungen in großem Stile sind“ (l. c.). So hat ja auch Sueß (s. oben) die Bezeichnung „Vorgräben“ durch „Vor-

tiefe“ ersetzt, schon weil sie seiner Meinung nach nicht ganz, nämlich hinsichtlich der Ausgestaltung ihres Innenrandes, dem Begriff des tektonischen Grabens entsprechen.

Die Bezeichnung „Vortiefe“ hat nun aber, wie wir gesehen haben, nur für einen Teil der abyssischen Rinnen, wenn auch für den überwiegenden, ihre Berechtigung, während andere dieser Rinnen „Rück“tiefen oder noch zweifelhaft hinsichtlich ihrer Lage zur Wirkungsrichtung des faltenden Druckes sind. Ihnen allen ist aber die Lage in der Umsäumung von Gebirgsbögen oder Inselzügen gemeinsam und sie alle sind in diesem Sinne abyssische „Saumtiefen“. Saumtiefe ist damit der Oberbegriff; viele Saumtiefen, aber nicht alle, so nicht diejenigen entlang der Westküste Südamerikas, sind zugleich auch Vortiefen.

III. Die Saumtiefen als Folge einer undatorischen Nahwirkung stabilerer Zonen.

In der Frage der Entstehung der abyssischen Tiefen erschien uns nicht ihre Lage vor den Gebirgs- und Inselzügen als wesentlich, — denn einzelne erfüllen garnicht die Forderung, Vortiefen zu sein, sondern sind Rücktiefen, während bei anderen Tiefen die Lage zur Faltungsrichtung noch zweifelhaft ist. Aber die Lage neben den Gebirgs- und Inselzügen scheint doch grundsätzlich wichtig zu sein, denn sonst wäre es schwer erklärlich, weshalb sich nicht auch fern von Gebirgs- und Inselzügen abyssische Gräben gebildet hätten[1]). —

Ich erblicke nun in den Tiefseegräben als den abyssischen „Saumtiefen“ einen besonders ausgeprägten Fall der weiter verbreiteten Erscheinung einer gesteigerten Senkung in der Umrandung von besonders stabilen Erdzonen. Eine „Saumsenke“, wenn gewiß auch nicht eine abyssische, ist in diesem Sinne auch die „norwegische Rinne“, die bei einer Tiefe von bis zu 700 Metern den Urgebirgssockel Norwegens von der flachen submarinen Nordseetafel trennt. Marcel Bertrand[2]) glaubte eine Deutung für diese

1) Die Romanche-Tiefe inmitten des Atlantik kommt hier, trotzdem sie abyssischer Art ist, kaum in Betracht, denn sie scheint schon morphologisch von den übrigen Gräben recht verschieden zu sein, — und vielleicht liegt ja auch sie, was ich mit Bezugnahme auf die nachfolgenden Ausführungen hier vorwegnehmen möchte, in der Randzone eines stabileren Erdstückes.

2) Marcel Bertrand, Sur la Continuité du phénomène de plissement etc. Soc. géol. France 3. Série, Band XX. 1892.

„bisher unerklärliche, eigenartige Erscheinung“ darin zu finden, daß er sie mit der oscillatorischen Bewegung der skandinavischen Kuppel in Verbindung brachte und dabei insbesondere annahm, daß der Raum der Senkung über den Raum der vorangegangenen Hebung hinausgegriffen habe und dadurch eben eine Randzone entstanden sei, die nur gesunken, aber nicht vorher gehoben wäre. F. Machatscheck[1]) erblickt in dieser Rinne einen tektonischen Graben. —

In den letzten Jahren hat mich viel die Frage beschäftigt, welche allgemeineren Gesetzmäßigkeiten hinsichtlich der ganzen Art der epirogenetischen Undation wohl bestehen könnten. Sie scheint mir von grundlegender Bedeutung für das Verständnis tektonischer Dinge und deshalb in hohem Maße würdig einer Verfolgung zu sein; denn ist es richtig, daß die Undation die nachfolgenden orogenetischen Vorgänge in hohem Maße bedingt, und zwar nicht nur hinsichtlich ihrer Intensität, was schon in dem Hall-Dana'schen Gesetze der Geosynklinalen zum Ausdruck kommt, sondern vielfach auch hinsichtlich ihrer ganzen Art, so kommen wir ja erst zu einem vertiefteren Verständnis der Orogenese, wenn wir in die Gesetze der sie bedingenden epirogenetischen Vorgänge etwas mehr Licht gebracht haben werden. Dann werden wir uns z. B. in der Frage des Eintretens reneganter Gebirgsbildung nicht mit der einfachen Annahme einer „Richtungsänderung des tektonischen Druckes“ zu begnügen brauchen, sondern wir werden vielleicht, indem wir die Vorbedingungen der Faltung zurückzuverfolgen suchen bis zu den frühesten Anlagen der nach der ersten Faltung in reneganter Richtung ausgebildeten Becken, verstehen lernen, weshalb notwendigerweise die Änderung der Faltungsrichtung als Folge jener Verhältnisse, die zu der Undation in neuer Richtung geführt haben, eintreten mußte. Und daß wir diese Frage gerade auf deutschem Boden zu fördern versuchen, ist m. E. deswegen gegeben, weil hier in dem Gegensatze der älteren (variscischen) und jüngeren (saxonischen) Tektonik ein ganz besonders auffälliger Fall reneganter Gebirgsbildung vorliegt. Das mag nur ein Beispiel für die Perspektiven sein, die uns die nähere Beschäftigung mit den Vorgängen der Undation hinsichtlich des Verständnisses der Tektonik eröffnet. So rechtfertigt sich der Versuch, einmal solchen Dingen nachzugehen. —

---

1) F. Machatscheck, Über epirogenetische Bewegungen. Penck-Festschrift, 1918, S. 9.

Ich habe früher[1]) ausgeführt, daß ich den epirogenetischen Vorgang des Einsinkens der Becken (Geosynklinalen) und des gleichzeitigen Aufsteigens der Schwellen (Geantiklinalen) als einen Wellenwurf größter Spannweite, als eine „Undation", glaube auffassen zu müssen, und zwar auf Grund dessen, daß diese vertikalen Bewegungen unter Verbiegung der von ihnen betroffenen großen Einheiten vor sich gehen. Die aufsteigenden Schwellen sind die Wellenberge, die einsinkenden Becken die Wellentäler. Von der orogenetischen Faltung, der „Undulation", ist dieser epirogenetische Wellenwurf zunächst durch seine weite Amplitude und dann durch seinen säkulären Fortgang unterschieden. Aber ein drittes scheint hinzuzukommen. —

Ob beim orogenetischen Faltenwurfe ein Erdstück zum Sattel oder zur Mulde wird, ist nicht durch seine Vorgeschichte bedingt, vielmehr sind die sich aufwölbenden und die sich einwölbenden Gesteinsmassen vor der Faltung von im allgemeinen gleicher Struktur gewesen. In dieser Hinsicht verhält es sich mit den Sätteln wie mit den „Schollenhorsten" und mit den Mulden wie mit den „Schollengräben". Doch anders ist die Sachlage bei der Undation. Hier scheint die innere Struktur dafür, ob ein Erdstück zu einem „Sattel" oder einer „Mulde" des Wellenwurfes größter Spannweite wird, oder, was dasselbe sagt, ob es in aufwärtige oder abwärtige säculäre Bewegung kommt, von großer Bedeutung zu sein. Jedenfalls ist, wenn wir die Bezirke der positiven und negativen epirogenetischen Bewegung vergleichend betrachten, in vielen Fällen unverkennbar, daß die aufsteigenden Bezirke (Schwellen) eine stabilere Struktur besitzen und auch von vornherein besessen haben, als die absteigenden (Becken), d. h., daß die stabileren Einheiten weit eher in den Zustand der Aufwärtsbewegung als die mobileren kommen, während umgekehrt die mobileren Einheiten im allgemeinen die sinkenden sind. Hierin scheint ein Grundgesetz aller epirogenetischen Vorgänge zu liegen, — so komplex diese ganzen Fragen auch sind und so viel auch Sonderumstände modifizierend hinzutreten und zu scheinbaren Ausnahmen führen. Damit zeigt aber das Undationsprofil Kurven von höchst unregelmäßiger Gestaltung. Gewisse Zonen der Erde haben sich lange im Zustande des Sinkens befunden, — werden dann aber unter Umständen zu Schwellengebieten, nachdem sie gefaltet worden und dadurch

---

1) Vgl. u. a. Tektonische Evolutionen und Revolutionen in der Erdrinde. Leipzig 1913, S. 17 ff.

und durch begleitende Intrusionen, — in manchen Fällen vielleicht auch durch außergewöhnlich mächtige Deckenergüsse, — versteift worden sind. Auch bei neu einsetzender Undation innerhalb der Faltungszonen scheinen gerade wieder die stabilsten Einheiten häufig einer neuen Absenkung zu widerstehen und unter Umständen in den Zustand von Spezialschwellen überzugehen, während die mobileren Zonen in erster Linie wieder in abwärtige Bewegung geraten. Sind Faltung, Intrusion und Aufwärtsbewegung des starreren Grundgebirges und unter Umständen auch mächtige Deckenergüsse die wichtigsten „versteifenden" Ereignisse, so ist umgekehrt die Senkung, der die Sedimentation folgt, ein den Untergrund stark mobilisierender Vorgang. Wir müssen davon ausgehen, daß wesentliche Mobilitätsunterschiede ja nur in einer verhältnismäßig dünnen äußeren Zone der Erdkruste anzunehmen sind, während sie sich nach der Tiefe mit der allgemeinen Mobilitätszunahme stark verschwächen und endlich wohl verschwinden müssen. Die ganze Erdschale von stark differentieller Mobilität besteht im Bereiche der „Massen" aus kristallinem oder doch stark gefaltetem und deswegen recht resistentem „Grundgebirge", in einer Geosynklinale aber wenigstens im oberen Teile aus leichter faltbarem Sedimente, während die Mächtigkeit des resistenteren „Grundgebirges" hier entsprechend verringert und damit die Gesamtresistenz um einen gewissen Teil schwächer ist; die Gesamtresistenz vermindert sich ja sozusagen mit jeder Menge von neuem Sediment, für die eine gleiche Menge der einsinkenden Grundgebirgsunterlage nach der Tiefe zu aus der differentiell mobilen Erdschale ausscheidet und in die Region der allgemein hohen Tiefenmobilität übertritt. So kommt es m. E., daß, wenn ein Gebiet einmal ins Sinken gekommen ist, dieser Zustand, da durch das Sinken ständig die Mobilität erhöht wird, auch gewahrt zu bleiben pflegt, bis eine neueinsetzende Orogenese nicht nur Aufwärtsbewegung, sondern auch gründliche Änderung in den Mobilitätsverhältnissen herbeiführt. Daß in den Übergangszonen zwischen Geosynklinalen und Geantiklinalen bald Hebung, bald Senkung und im Zusammenhange damit bald Regression, bald Transgression eintreten kann, ist eine Sache für sich, die mit Verlegungen des Bewegungsscharniers zwischen der aufsteigenden und der sinkenden Region und mit gewissen Veränderungen in der Amplitude des undatorischen Wellenwurfes zusammenhängt, was vielleicht durch Veränderungen in der Stärke des undatorischen Druckes bedingt ist. —

Aber nicht allein die Struktur scheint für die aufwärtige

oder abwärtige Richtung und für den Betrag der vertikalen Bewegung wesentlich zu sein, sondern auch gewisse Lageverhältnisse scheinen in Betracht zu kommen. Es scheint nämlich bei der epirogenetischen Hebung und Senkung nicht nur auf die absolute Mobilität, sondern auch auf die relative Mobilität, d. h. auf die Mobilität im Vergleich mit derjenigen der Nachbarzonen, anzukommen. Ein bestimmter Mobilitätsgrad kann einen Bezirk zu einer aufsteigenden Antiklinale machen, wenn nämlich dieser Bezirk zwischen Gebieten weit geringerer Mobilität liegt, während ein Bezirk gleicher Mobilität neben oder zwischen stabileren Einheiten zur sinkenden Zone werden kann. Bei gleicher absoluter Mobilität ist aber die relative Mobilität einer Örtlichkeit um so größer, je mehr sie der starreren Einheit genähert ist. Allerdings nimmt ja, — das liegt im Wesen des undatorischen Wellenwurfes, wie er z. B. in der Bildung der Geosynklinalen sich ausdrückt, begründet, — mit der Annäherung an die Schwellen auch die Mächtigkeit der mobilen Schichtmassen der geosynklinalen Region, d. h. die absolute Mobilität, ab, und es können sich dann die Abnahme der absoluten Mobilität und die Zunahme der relativen Mobilität bis zu einem gewissen Grade aufheben. —

Haben wir somit im allgemeinen in den mobileren Zonen eine epirogenetische Senkung, so haben wir häufig innerhalb dieser sinkenden Zonen eine gesteigerte Senkung in der Nachbarschaft der stabileren Region, und wohl besonders dann, wenn stabile und mobile Einheiten einigermaßen unvermittelt aneinander grenzen. Das läßt sich in manchem Falle aus den Sedimentationsverhältnissen der Vorzeit, insonderheit aus der ersten Anlage neuer Becken und aus den Mächtigkeitsverhältnissen der Sedimente, herauslesen. — Man könnte hier geradezu von einer Art „Nahwirkung“ der stabileren Einheiten auf ihre nächste Umgebung sprechen, wie eine solche auch bei der Orogenese vorhanden ist. Was diese anlangt, so habe ich an anderer Stelle[1]) ausgeführt, daß für Art und Intensität der Faltung nicht nur die Mobilität des Untergrundes, d. h. die „Gefügigkeit gegen den orogenetischen Druck“, sondern auch die Lage zu den Schwellengebieten, d. h. die „Erreichbarkeit für den orogenetischen Druck“[2]),

---

1) H. Stille, Injektivfaltungen und damit zusammenhängende Erscheinungen Geol. Rundschau Bd. VIII, 1917, S. 89 ff., insbes. S. 101 ff.

2) Dabei vertrete ich, wie auch aus meinen früheren Veröffentlichungen ersichtlich ist, die Auffassung, daß der faltende Druck nicht in der Richtung von der Geosynklinale auf die Schwellen, sondern umgekehrt in der Richtung von den Schwellen auf die Geosynklinalen wirkt und daß die Einseitigkeit der

von Bedeutung ist. Haben wir doch in weiten Geosynklinalen in großen Abständen von der Schwelle eine gewiß nicht geringere Mobilität, als in den schwellennäheren Zonen, und doch bleiben diese schwellenferneren Gebiete oft ungefaltet, während in den schwellennahen Zonen intensive Faltung erfolgt. Und wie bei der Orogenese, so haben wir also in manchen Fällen, wie es scheint, auch bei der Epirogenese besondere Nahwirkungen der stabileren Einheiten, und mit solchen bringe ich die Tiefseegräben in Zusammenhang. Hier sind die Gebirgs- oder Inselbögen die Zonen höherer Stabilität, und zwar haben sie diese wohl nicht ausschließlich durch Faltung, sondern, — und in manchem Falle vielleicht in besonderem Maße, — auch durch Intrusion eruptiver Gesteinsmassen erhalten. Die orogenetische Nahwirkung der Schwellengebiete führt also zu den Gebirgskränzen, die epirogenetische Nahwirkung zu den „Saumtiefen" und im Falle geringer Zuschüttung dieser zu der morphologisch besonders gekennzeichneten Form der Tiefseegräben. Die epirogenetische Nahwirkung bereitet aber, indem sie besonders starke Mobilität in besonders schwellennahe Zonen trägt, die Stätte für eine gesteigerte orogenetische Nahwirkung vor. —

Man wird vielleicht fragen, warum die Erscheinung der Tiefseegräben, wenn sie durch das Nebeneinander der stabileren und mobileren Einheiten bedingt erscheint, nicht viel allgemeiner auftritt, als der Fall ist. In dieser Hinsicht verweise ich zunächst auf die früheren Ausführungen (vgl. Seite 346), die dahin gingen, daß die tektonischen R a n d s e n k u n g e n sicher viel verbreiteter sind, als die morphologischen Rand t i e f e n, nur sind die Senkungen aufgefüllt und dann nicht mehr erkennbar. Aber zweitens kommt es sicher auch auf die Art des Überganges von der Schwelle zur Senke an, vor allem wohl auf ein einigermaßen u n v e r m i t t e l t e s Nebeneinander der stabileren und mobileren Einheiten, wie ja auch bei der Orogenese wohl die Beschaffenheit der Übergangszone die Erklärung dafür geben kann, daß nicht überall die Festlandsschwellen sich mit Gebirgskränzen umziehen, sondern daß die Kränze lückenhaft sein können oder daß wenigstens die Stärke der Faltung und die Breite der Faltenzone in der Umrandung derselben Schwelle sehr verschieden sind, trotzdem auch dort, wo der Faltenkranz unterbrochen oder

---

Faltung in der Richtung auf die Schwellen das Ergebnis einer Vorwärtsbewegung und eines Ü b e r q u e l l e n s der bei der Faltung aufsteigenden Füllung der Geosynklinalen in der Richtung auf den Rand der starren Schwellen und über diesen hinweg ist.

schwach entwickelt ist, die Schwelle von der Geosynklinale umgeben war.

Wie würde sich nun endlich aus den Mobilitätsverhältnissen und deren Veränderungen die Erscheinung der fossilen Vortiefen und das damit zusammenhängende Wandern der Faltung erklären? —

Wir fassen gewisse Übergangszonen zwischen einer Festlandsschwelle und einer Geosynklinale ins Auge, in denen sich der Strand mehrfach hin und her schiebt und in denen es wenigstens zeitweilig zu Sedimentationen kommt, die aber erstens an sich schon geringer sind, als die gleichzeitigen Sedimentationen in der eigentlichen Geosynklinale, und die zweitens vielleicht auch in einer eingeschalteten Hebungszeit wenigstens teilweise wieder beseitigt werden. Somit unterscheidet sich eine derartige Übergangszone von der eigentlichen Schwelle durch einen gewissen Absatz von Sediment, von der eigentlichen Geosynklinale durch die geringere Mächtigkeit dieses Sedimentes und sie ist damit wohl von etwas höherer Mobilität als die Festlandsschwelle, aber natürlich von weit geringerer Mobilität als die Geosynklinale. Setzt nun eine Faltung ein, so wirkt diese wegen der stärkeren Resistenz des Bodens hier nicht oder doch sehr viel schwächer, als in der eigentlichen Geosynklinale. Solche ungefaltet bleibende Gebiete mit verhältnismäßig geringer Bedeckung durch jüngeres Sediment finden sich häufig, und zwar oft in großer Breite, in den Randzonen der Schwellen, und ich möchte sie als deren „Höfe“ oder „Hofzonen“ bezeichnen. Ich erinnere an den Kanadischen Schild; das alte Grundgebirge verschwindet hier nach Südosten unter flachliegenden und kaum gestörten Tafeln paläozoischen Sedimentes und erst weiterhin setzt der jungpaläozoische Faltenwurf ein, d. h. erst weiterhin lag der eigentliche „Trog“ der Alleghanies. Oder ich verweise auf den Baltischen Schild; in seinem Süden ist in Schonen oder Bornholm die palaeozoische Schichtfolge bei verhältnismäßig geringer Mächtigkeit zunächst noch ungefaltet, und erst weiterhin folgt das tiefere Becken, aus dem die variscischen Alpen emporgestiegen sind. —

In solchen Hofzonen der alten Massen entstehen, nachdem die angrenzende Geosynklinale aufgefaltet worden ist, die Vortiefen. Daß die Hofzone, ehemals ein Übergangsgebiet zwischen Masse und Geosynklinale, vor dem nunmehr entstandenen Faltenbogen, wie wir aus der nach der Faltung einsetzenden starken Sedimentation herauslesen, ins Sinken gerät, glaube ich dadurch erklären zu können, daß, während sie vorher hinsichtlich ihrer Mo-

bilität eine Zwischenstellung zwischen einer stabileren Einheit, der Schwelle, und einer mobileren, der Geosynklinale, einnahm, sie nach erfolgter Auffaltung und Versteifung der Geosynklinalzone zu beiden Seiten von stabileren Erdzonen umgrenzt wird, also eine hohe relative Mobilität besitzt.

Aber auch dann, wenn die künftige Vorzone nicht als „Hofzone" vorgebildet war, sondern wenn Schwelle und Geosynklinale ohne vermittelnde Zwischenzone aneinandergrenzten, mag unter Umständen schon die durch Faltung und Intrusion erzielte Stabilität des neuen Gebirgsbogens genügt haben, um seine unmittelbare Vorzone durch „Nahwirkung" (vgl. oben) zum Sinken zu bringen. —

So tritt in der Hofzone eine gesteigerte Senkung und als Folge derselben bei der nächsten orogenetischen Phase die Faltung ein, und zwar haben wir dann ein Wandern der Faltung in der Richtung auf die Festlandsschwellen, das im Gegensatz steht zu den Wanderungen in der Richtung auf die Geosyklinale, wie sie sich bei der variscischen Faltung gegenüber der kaledonischen und bei der alpiden Faltung gegenüber der variscischen vollzogen haben. Insgesamt ergibt sich für den Boden Europas ein eigenartiger Rhytmus des Wanderns, ein zweimaliges starkes Vorwandern nach Süden (geosynklinalwärts), jeweilig gefolgt in den Unterphasen der Faltungen von kleineren Rückwanderungen nordwärts (rahmenwärts), die durch Vortiefenbildung vorbereitet waren.

Über analytische Berechnung der Geschosspendelungen.

Von

F. Noether[1]) in Karlsruhe i. B.

Vorgelegt von C. Runge in der Sitzung vom 30. Mai 1919.

Das allgemeine ballistische Problem, insofern es nicht nur die Bestimmung der als eben angenommenen Schwerpunktsbahn, sondern auch die Seitenabweichungen aus der Schußebene und die Relativbewegung des Langgeschosses zum Schwerpunkt, die Pendelung, umfaßt, ist nur in begrenztem Umfang behandelt; sei es durch Beschränkung auf nur qualitative Diskussion[2]), sei es, quantitativ, durch Voraussetzung sehr flacher Bahnen[3]) oder durch Verwendung graphischer Methoden[3]), die die Zugrundelegung spezieller Zahlenwerte fordern. Die ballistische Praxis hat aber auf Fragen geführt, die eine quantitative, analytische Durchführung dieses Problems in weitergehendem Maß erfordern. Das sind insbesondere die folgenden Fragen: Welchen Gesetzen gehorcht die Seitenabweichung der Geschosse? Wieviel Pendelungen macht das Geschoß längs der Bahn? Von welcher Größenverbindung hängt die Zahl der Pendelungen ab? Wann ist ein Geschoß steif (Bodentreffer), wann ist es folgsam (Spitzentreffer)? Wie groß sind die Pendelungs-

1) Mit Benützung eines unveröffentlichten Manuskripts von A. Sommerfeld. In demselben waren die Begriffe des „steifen" und des „absolut folgsamen" Geschosses eingeführt und die Näherungslösung des § 2 (jedoch in einer Form, die ihren Gültigkeitsbereich nicht erkennen ließ und unter spezielleren Voraussetzungen) sowie die daraus folgende Berechnung der Seitenabweichung aufgestellt.

2) Klein-Sommerfeld, Theorie des Kreisels. IV, S. 931 f.

3) Vgl. insbesondere: C. Cranz, Ztschr. f. Math. u. Phys. 43 (1898) und Lehrb. d. Ballistik, Bd. 1 (Leipzig 1917), Nr. 51—57 u. d. zugehörigen Literaturnoten.

amplituden und in welchem Maß wird der Widerstand des Geschosses durch die Pendelungen erhöht? Die letzte Frage muß zur Erklärung der Abweichung der Geschoßbahnen von den „ballistischen Kurven" gestellt werden, deren Ableitung bekanntlich voraussetzt, daß der Geschoßwiderstand nur durch die Form und momentane Geschwindigkeit des Geschosses bestimmt sei.

Um einen solchen analytischen Ansatz zu begründen, werden im Folgenden einige vereinfachende Voraussetzungen gemacht:

I. Die Schwerpunktsbahn wird in erster Näherung als bekannte ballistische Kurve vorausgesetzt und zwar reicht ihre Darstellung durch die Hodographengleichung (d. i. die Beziehung zwischen der Größe und der Richtung der Geschwindigkeit) aus. Da wir uns auf Anfangsgeschwindigkeiten beschränken, für die das quadratische Widerstandsgesetz gilt[1]) — die Verzögerung sei

$$q = cv^2 \tag{1}$$

— ist diese Gleichung in unserem Falle durch Quadraturen integrabel[2]).

II. Außer dem Widerstand wirkt auf das Geschoß ein Widerstandsmoment, das die Geschoßaxe von der Bahntangente abzudrehen, und zu ihr quer zu stellen strebt, sodaß die tangentiale Stellung eine labile Gleichgewichtslage darstellt. Sein Momentvektor steht senkrecht auf der Bahntangente und der Geschoßaxe, es verschwindet bei tangentialer Stellung (Anstellwinkel $\vartheta = 0$) und bei annähernder Querstellung ($\vartheta = \pi/2$). Da es im oben angegebenen Geschwindigkeitsbereich gleichfalls mit v^2 proportional ist, wurde von verschiedenen Seiten[3]) der Ansatz

$$M = Cv^2 \sin\vartheta \cos\vartheta$$

empfohlen. Wir beschränken die Untersuchung auf so kleine Ausschlagswinkel (etwa $\vartheta < 30^0$), daß dieser Ansatz noch zu

$$M = Cv^2 \sin\vartheta \,^{4)} \tag{2}$$

vereinfacht werden kann. In dieser Form ist er empirisch gut bestätigt.

---

1) Das ist bekanntlich bis in die Nähe der Schallgeschwindigkeit der Fall. Die am meisten interessierenden, in der Praxis vorkommenden Steilbahnen fallen in dieses Gebiet.

2) Cranz, Lehrb. d. Ballistik, Bd. 1, S. 118.

3) Klein-Sommerfeld, IV, S. 932. Siehe auch R. Grammel: Zusätze zur Kreiseltheorie, Ztschr. f. Math. u. Phys. (64) (1916), § 3f.

4) Dieser Ansatz findet sich auch bei N. Sabudski: Bewegung der Langgeschosse, Stuttgart 1906.

III. Durch den dem Geschoß erteilten Drehimpuls um die Längsaxe (N) wird die eben bezeichnete labile Gleichgewichtslage in eine stabile verwandelt, wenn

$$N^2 > 4ACv^2 \qquad (3)$$ [1]

(A = Querträgheitsmoment für den Schwerpunkt). Die momentane Geschoßbewegung wird dann eine Präzession um die (fest gedachte) Bahntangente, überlagert von einer Nutation. Falls die Ungleichung (3) hinreichend stark erfüllt ist, wird die Nutationsamplitude klein neben der der Präzession (pseudoreguläre Präzession), ausgenommen den (hier auszuschließenden) Fall, daß durch starke seitliche Stöße schon zu Beginn der Bewegung „Stoßnutationen" erzeugt werden. Wir setzen daher die Momentanbewegung der Geschoßaxe als reine Präzession voraus, bei der die Impulsaxe mit der Geschoßaxe zusammenfällt.

§ 1. Der analytische Ansatz.

Die Grundgleichung für die Bewegung des Geschosses relativ zu seinem Schwerpunkt ist die Impulsgleichung:

$$\frac{d\overline{N}}{dt} = \overline{M} \qquad (4)$$ [2]

wobei $\overline{N}$ den Impulsvektor, $\overline{M}$ den Momentvektor darstellt, und zwar ist dies die Form der Gleichung für ein, im übrigen beliebiges, richtungsfestes Koordinatensystem im Schwerpunkt. Für die vorliegende Aufgabe kommt mehr die Beziehung auf ein geeignetes, mit dem Geschoß mitbewegtes Koordinatensystem mit veränderlichen Axenrichtungen in Betracht, also die Gleichung:

$$\frac{d\overline{N}}{dt} + \overline{\omega N} = \overline{M}, \qquad (4')$$

wobei $\overline{\omega}$ den Drehungsvektor des (rechtssinnigen) Koordinatensystems bedeutet. Als bewegtes Koordinatensystem wählen wir die Normale, Tangente und Binormale der zu Grunde gelegten

---

1) Vgl. Klein-Sommerfeld, II, S. 323. An Stelle der dort mit P bezeichneten Größe tritt hier das Moment Cv^2.

2) Die Schreibweise der Vektoren ist die Résal'sche (durch übergesetzten horizontalen Strich, vgl. Heun, Kinematik, Leipzig 1906, S. 8ff.). Das skalare Produkt zweier Vektoren wird durch Nebeneinandersetzen, $\overline{N}\,\overline{v}$, bezeichnet, das vektorielle durch Nebeneinandersetzen mit gemeinsamem Strich: $\overline{Nv}$.

ebenen ballistischen Kurve, für das die Impulskomponenten seien: S, T, Z. Der Drehungsvektor $\overline{\omega}$ ist für dieses System: $0, 0, -d\varphi/dt$, wenn φ der Winkel zwischen Tangente und Horizontaler ist. Der Momentvektor ist nach Voraussetzung II und III:

$$\overline{M} = \frac{Cv}{N}\,\overline{Nv} \tag{2'}$$

und hat nach unseren Axen die Komponenten

$$-\frac{Cv^2 Z}{N},\ 0,\ \frac{Cv^2 S}{N}.$$

Also haben wir nach (4'):

$$\begin{aligned} \frac{dS}{dt} &= -\frac{d\varphi}{dt}T - \frac{Cv^2 Z}{N} \\ \frac{dT}{dt} &= \frac{d\varphi}{dt}S \\ \frac{dZ}{dt} &= \frac{Cv^2 S}{N}. \end{aligned} \tag{4''}$$

Ersichtlich folgt aus diesem System:

$$S^2 + T^2 + Z^2 = N^2 = \text{const.}$$

sodaß es sich um ein lineares Gleichungssystem handelt. In diese Gleichungen führen wir durch Multiplikation mit $-\cos^2\varphi\, dt/d\varphi$ als unabhängige Variable an Stelle der Zeitvariablen t die rein geometrische und unbenannte, ständig wachsende Größe:

$$u = -\operatorname{tg}\varphi \tag{5}$$

ein. Da für die ballistische Kurve die Normalbeschleunigung $-g\cos\varphi$ besträgt, ist:

$$-v\frac{d\varphi}{dt} = g\cos\varphi$$

und man erhält:

$$\begin{aligned} \frac{dS}{du} &= \frac{T}{1+u^2} - (1+u^2)\frac{Cv^3\cos^3\varphi}{Ng}Z \\ \frac{dT}{du} &= -\frac{S}{1+u^2} \\ \frac{dZ}{du} &= (1+u^2)\frac{Cv^3\cos^3\varphi}{Ng}S. \end{aligned} \tag{6}$$

In diesem System kommt außer der unabhängigen Variablen u noch die Größe $v\cos\varphi$ vor. Diese Geschwindigkeitskomponente

kann aber mittels der Hodographengleichung der ballistischen Kurve durch φ, somit auch durch u, ausgedrückt werden. Durch Elimination der Zeitvariabeln t folgt nämlich aus den ballistischen Grundgleichungen (vgl. z. B. Cranz, Lehrb. d. Ballistik, I, Nr. 17)

$$\frac{d(v\cos\varphi)}{d\varphi} = \frac{cvf(v)}{g} \tag{7}$$

worin $cf(v)$ die allgemeine Widerstandsfunktion bedeutet. Nach Voraussetzung I wählen wir $f(v) = v^2$. Die Gleichung (7) ist in dem Fall integrabel und ergibt mit der unabhängigen Variabeln u:

$$\left(\frac{v_s}{v\cos\varphi}\right)^2 = 1 + \frac{cv_s^2}{g}\left(\mathfrak{Ar}\,\mathfrak{Sin}\,u + u\sqrt{1+u^2}\right) = 1 + c'\chi(u), \tag{7'}$$

wenn noch v_s die Scheitelgeschwindigkeit (für $u = 0$) und c' und $\chi(u)$ die aus (7′) zu ersehenden Abkürzungen bezeichnen. c' und auch $c'\chi(u)$ ist in den in der Praxis vorkommenden Fällen, wenigstens bei Steilbahnen, immer eine kleine Zahl. Die Annahme $c' = 0$ würde bedeuten, daß als ballistische Kurve eine widerstandsfreie Parabel zu Grunde gelegt ist.

Mittels (7′) haben wir also in (6)

$$\frac{Cv^3\cos^3\varphi}{Ng} = \lambda\left(1 + c'\chi(u)\right)^{-\frac{3}{2}}$$

worin

$$\lambda = \frac{Cv_s^3}{Ng} \tag{8}$$

eine charakteristische Konstante unseres Problems bezeichnet. Sie kann (vgl. Cranz, Lehrb. d. Ballistik, I, S. 372) als das Verhältnis der Winkelgeschwindigkeiten der Präzessionsbewegung einerseits, der Richtungsänderung der Bahntangente andererseits, und zwar im Scheitel ($u = 0$) der Bahn gedeutet werden. Wir führen nun, um die Gleichungen (6) noch übersichtlicher zu gestalten, vorübergehend eine neue unabhängige Variable ψ ein durch die Definition:

$$\psi = \int_0^u (1+u^2)\left(1 + c'\chi(u)\right)^{-\frac{3}{2}} du. \tag{9}$$

Sie nehmen dann die einfache Form an:

$$\begin{aligned} \frac{dS}{d\psi} &= \gamma(u) T - \lambda Z \\ \frac{dT}{d\psi} &= -\gamma(u) S \qquad (6') \\ \frac{dZ}{d\psi} &= \lambda S \end{aligned}$$

$$\text{mit } \gamma(u) = (1+u^2)^{-2}(1+c'\chi(u))^{\frac{3}{2}}.$$

§ 2. Flachbahnen, steifes und folgsames Geschoss.

Die Gleichungen (6′) legen es nahe, zwei charakteristische Grenzfälle voranzustellen:

1) Starkes Überwiegen der Glieder mit dem Faktor $\gamma(u)$ $(1 = \gamma(0) \gg \lambda)$: Vernachlässigung der Glieder mit λ kommt ersichtlich darauf hinaus, daß man in der ursprünglichen Gleichung (4) $\overline{M} = 0$ setzt, also $d\overline{N}/dt = 0$ hat: Die Impuls- (und damit die Geschoßaxe) bleibt sich also dauernd nahezu parallel; das Geschoß ist wohl stabil, aber nicht in dem gewünschten Sinn, daß seine Spitze möglichst der Bahntangente folgt. Es ist vielmehr „steif" und schlägt mit dem Boden im Ziel auf.

2) Starkes Überwiegen der Glieder mit dem Faktor λ $(1 = \gamma(0) \ll \lambda)$: Vernachlässigung der Glieder mit dem Faktor $\gamma(u)$ ist damit gleichwertig, daß man in den Gleichungen (4″) die Krümmung der Bahn $(d\varphi/dt)$ vernachlässigt. Die Bewegung wäre dann eine reine Präzession mit der durch anfänglichen Stoß bestimmten, im allgemeinen kleinen Amplitude. Man muß fragen, ob auch bei Beachtung der im Ganzen endlichen Richtungsänderung der Bahntangente die Präzessionsamplitude klein, das Geschoß also f o l g s a m bleibt, oder ob etwa durch die Richtungsänderung immer eine bestimmte Amplitude hervorgebracht wird. Zur Orientierung machen wir zunächst die Voraussetzung, daß $\gamma(u)$ als Konstante betrachtet werden könne (für flache Bahnen, wo $u = -\operatorname{tg}\varphi \approx 0$, ist $\gamma(u) \approx 1$. Dann folgt aus der zweiten und dritten Gleichung von (6′):

$$\lambda \frac{dT}{d\psi} + \gamma \frac{dZ}{d\psi} = \frac{d}{d\psi}(\lambda T + \gamma Z) = 0$$

also

$$R = \lambda T + \gamma Z = R_0$$

mit einer Konstanten R_0. Hiermit können wir T aus (6′) eliminieren und erhalten

$$\frac{dS}{d\psi} = \frac{\gamma}{\lambda} R_0 - \lambda Z (1 + \gamma^2/\lambda^2)$$
$$\frac{dZ}{d\psi} = \lambda S.$$

Die Integration ergibt

$$S = r_1 \cos \sqrt{\lambda^2 + \gamma^2}\, \psi + r_2 \sin \sqrt{\lambda^2 + \gamma^2}\, \psi$$
$$Z = \frac{\gamma R_0}{\lambda^2 + \gamma^2} + r_1' \cos \sqrt{\lambda^2 + \gamma^2}\, \psi + r_2' \sin \sqrt{\lambda^2 + \gamma^2}\, \psi$$
$$-\sqrt{1 + \gamma^2/\lambda^2}\, r_1' = r_2; \quad \sqrt{1 + \gamma^2/\lambda^2}\, r_2' = r_1.$$

Man erkennt also: Die Bewegung des Geschosses ist eine harmonische Pendelung, mit der Frequenz $\sqrt{\lambda^2 + \gamma^2}/2\pi$, um eine mittlere, nach rechts schräggestellte Lage. Dabei beträgt die mittlere Schrägstellung

$$\frac{Z}{N} = \frac{\gamma}{\lambda^2 + \gamma^2} \frac{R_0}{N}.$$

Verließ das Geschoß genau zentrisch das Rohr ($Z_0 = 0$), so ist $R_0 = \lambda T_0 = \lambda N$, also die mittlere Schrägstellung

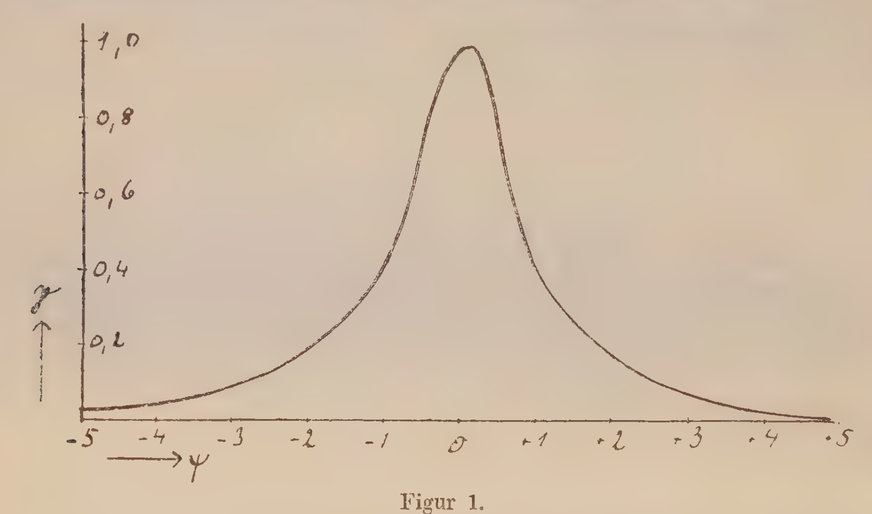
$$\frac{Z}{N} = \frac{\lambda \gamma}{\lambda^2 + \gamma^2}.$$

In diesem Falle würde, wegen der Periodizität der Lösungen 10), die Geschoßaxe nach Verlauf einer vollen Pendelung immer wieder genau zur Bahntangente zurückkehren und es wird

$$\text{(10′)} \quad \begin{aligned} \frac{S}{N} &= \frac{\gamma}{\sqrt{\lambda^2 + \gamma^2}} \sin \sqrt{\lambda^2 + \gamma^2}\, (\psi - \psi_0) \\ \frac{Z}{N} &= \frac{\gamma}{\sqrt{\lambda^2 + \gamma^2}} \left[1 - \cos \sqrt{\lambda^2 + \gamma^2}\, (\psi - \psi_0)\right]. \end{aligned}$$

worin ψ_0 den Anfangswert dieser Variabeln (für $t = 0$) bedeutet. Das Resultat dieser orientierenden Rechnung ist: Pendelung und mittlere Schrägstellung werden um so kleiner, je größer λ/γ ist; im Grenzfall $\lambda/\gamma \gg 1$ werden beide zu Null werden, das Geschoß wird also absolut folgsam und trifft mit der Spitze voran im Ziel auf. Wegen der vorausgesetzten Konstanz von γ ist die Pendelungsamplitude und die mittlere Schrägstellung konstant. Zeichnet man daher die Bahnkurve der Axe in einem mit dem

Schwerpunkt mit bewegten System auf einer Einheitskugel auf, so wird sie eine zykloidenartige Kurve und zwar eine gemeine Zykloide, in der allerdings die Höhe im Verhältnis $1/\sqrt{1+\gamma^2/\lambda^2}$ gegen die Länge verkürzt ist. War das Geschoß beim Verlassen des Rohres schon schräg gestellt, so geht, nach Gleichung (10), die gemeine in eine gestreckte oder verschlungene Zykloide über.



Figur 1.

Für den allgemeinen Fall nicht sehr flacher Bahnen muß die Veränderlichkeit von γ in Betracht gezogen werden. Wir fassen jetzt zweckmäßig die Komponenten S und Z zu ihrer komplexen Verbindung $S + iZ = Q$ zusammen und erhalten aus der ersten und dritten Gleichung von (6′)

$$\frac{dQ}{d\psi} - i\lambda Q = \gamma T, \tag{11}$$

wozu noch

$$T^2 = N^2 - S^2 - Z^2 \tag{11'}$$

tritt.

Die Integration von (11) ergibt nun:

$$Q = e^{i\lambda\psi} \int^{\psi} \gamma T e^{-i\lambda\psi}\, d\psi. \tag{12}$$

Wegen der Voraussetzung (II), daß die Anstellwinkel ϑ klein bleiben sollen (etwa $< 30^0$), und wegen $\cos\vartheta = T/N$ ist es der erste Schritt einer suczessiven Näherung, (dessen Fehler $< 13\,\%$), wenn man in (12) statt der noch unbekannten Größe T den konstanten Betrag des Impulses N einsetzt. Also

(12′)
$$Q = Ne^{i\lambda\psi} \int^{\psi} \gamma e^{-i\lambda\psi} \, d\psi.$$

Verläßt das Geschoß genau zentrisch das Rohr, so ist zu Anfang (ψ_0): $S = Z = Q = 0$, also

(12″)
$$Q = Ne^{i\lambda\psi} \int_{\psi_0}^{\psi} \gamma e^{-i\lambda\psi} \, d\psi.$$

Im allgemeinen Fall überlagert sich dieser Bewegungsform eine durch den Anfangsstoß bestimmte Zirkularschwingung:

$$Q_1 = Be^{i\lambda\psi}.$$

Man hat so den Satz: Durch die Geschütz- und Geschoßabmessungen ist der Verlauf der Geschoßbewegung im Wesentlichen schon gegeben. Die beim Abschuß entstehenden, zufälligen Störungen äußern sich nur in einer hinzutretenden zirkularen Pendelung von konstant (also im Allgemeinen klein) bleibender Amplitude. Hierher gehören z. B. die durch Rohrerweiterung oder durch gleichmäßigen Wind erzeugten Pendelungen. Nicht enthalten sind dagegen in diesen Schwingungen die unter Umständen beim Abschuß entstehenden „Stoßnutationen". Sie sind durch die Annahme III, daß die Impulsaxe mit der Geschoßaxe zusammenfallen sollte, ausgeschaltet.

Im weiteren lassen wir auch solche zufälligen Pendelungen beiseite, behandeln also nur die partikuläre Lösung (12″). Diese läßt sich formal in eine nach Potenzen von $1/\lambda$ fortschreitende Reihe entwickeln. Durch partielle Integration erhält man nämlich:

(13)
$$\begin{aligned} \frac{Q}{N} &= e^{i\lambda\psi} \int_{\psi_0}^{\psi} \gamma e^{-i\lambda\psi} \, d\psi \\ &= e^{i\lambda\psi} \left[\frac{i}{\lambda} \left(\gamma e^{-i\lambda\psi} - \gamma_0 e^{-i\lambda\psi_0} \right) - \frac{i}{\lambda} \int_{\psi_0}^{\psi} \gamma' e^{-i\lambda\psi} \, d\psi \right] \end{aligned}$$

(mit $\gamma' = d\gamma/d\psi$). Der vom Integralzeichen freie Teil dieser Lösung lautet, in reellen und imaginären Teil zerlegt:

(13′)
$$\begin{aligned} \frac{S^{(1)}}{N} &= \frac{\gamma_0}{\lambda} \sin \lambda (\psi - \psi_0) \\ \frac{Z^{(1)}}{N} &= \frac{\gamma}{\lambda} - \frac{\gamma_0}{\lambda} \cos \lambda (\psi - \psi_0). \end{aligned}$$

Es erscheint also hier die früher für Flachbahnen unter der Voraussetzung konstanter γ bereits abgeleitete Form (10′) als erstes Entwicklungsglied, mit der Modifikation, daß λ an Stelle von $\sqrt{\lambda^2+\gamma^2}$ getreten ist, die darauf hindeutet, daß die hier benutzte Voraussetzung kleiner Anstellwinkel ($T \approx N$) nur im Falle hinreichend großer λ/γ realisiert ist. Ferner besteht der Unterschied, daß das erste Glied der zweiten Gleichung (13), die mittlere Schrägstellung γ/λ, jetzt als veränderlich zu betrachten ist: Wir haben wieder eine zykloidenartige Bewegung, wie bei Flachbahnen, aber nur zu Beginn und am Ende der Bewegung den Charakter der gemeinen Zykloide, in der Gegend des Scheitels dagegen der gestreckten (da $\gamma \geqq \gamma_0$).

Unter welchen Umständen wird dieses Glied die wirkliche Bewegung, insbesondere die Pendelungen, schon richtig darstellen? Dazu ist erforderlich, daß das mit dem Integralzeichen behaftete Restglied in (13) klein bleibt neben γ_0/λ. Der Integrand dieses Restgliedes zerfällt in den periodischen Faktor $e^{-i\lambda\psi}$ und den Faktor γ'. Wegen der Veränderlichkeit von γ' wird das Integral nicht rein harmonisch periodisch, sondern wird einen gewissen, von 0 verschiedenen übrigens veränderlichen Mittelwert erhalten, der zur Pendelungsamplitude beiträgt. Insbesondere ist wesentlich, daß γ' bei $\psi = 0$ einen raschen Anstieg, dann Abfall von positiven zu negativen Werten hat, sodaß sich Bestandteile in dem Integral addieren, die bei konstantem γ' sich gegenseitig aufheben würden. Diese Teile werden, wegen der Oszillationen des ersten Faktors, von der Größenordnung des Maximalwertes von γ'/λ, zur Pendelungsamplitude kommen also im absteigenden Ast ($\psi > 0$) Beträge von der Ordnung γ'/λ^2. Da der Maximalbetrag von γ' endlich bleibt, so kann λ sicher so groß gewählt werden, daß diese Glieder klein werden gegen γ_0/λ. Für sehr große λ wird nun die mittlere Schrägstellung und die Pendelungsamplitude nach (13′) sehr klein. Man findet also:

Für sehr große λ wird (wie für Flachbahnen), die Bewegung der Geschosse durch das erste Glied von (13), dargestellt. Die Geschoßaxe folgt in dem Fall mit geringer Schrägstellung nach rechts und noch kleinerer Pendelungsamplitude, der Bahntangente. **Für den Grenzfall $\lambda = \infty$ wird das Geschoß absolut folgsam.**

Durch nochmalige partielle Integration erhält man aus (13):

$$\frac{Q}{N} = \frac{i}{\lambda}\left(\gamma - \gamma_0 e^{i\lambda(\psi-\psi_0)}\right) + \frac{1}{\lambda^2}\left(\gamma' - \gamma_0' e^{i\lambda(\psi-\psi_0)}\right) - \frac{e^{i\lambda\psi}}{\lambda^2}\int_{\psi\gamma}^{\psi} {}'' e^{-i\lambda\psi}\,d\psi. \tag{14}$$

Für das mit dem Integralzeichen behaftete Restglied in (14) gilt analoge Diskussion wie oben für das in (13). Wenn es (für genügend große λ) vernachlässigt wird, werden die Glieder zweiter Ordnung:

$$(14') \qquad \begin{aligned} \frac{S^{(2)}}{N} &= \frac{\gamma'}{\lambda^2} - \frac{\gamma_0'}{\lambda^2} \cos \lambda (\psi - \psi_0) \\ \frac{Z^{(2)}}{N} &= \qquad \frac{\gamma_0'}{\lambda^2} \sin \lambda (\psi - \psi_0), \end{aligned}$$

also wieder eine veränderliche, mittlere Schrägstellung der Axe, und zwar hier in der Schußebene, nach oben im aufsteigenden, nach unten im absteigenden Ast der Flugbahn, überlagert wieder von einer zirkularen Pendelung.

Die hier begonnene Reihenentwicklung nach Potenzen von $1/\lambda$ könnte durch fortgesetzte partielle Integration fortgeführt werden, würde sich aber nur als s e m i k o n v e r g e n t ergeben. Das erkennt man, wenn man z. B. die sich für die Schrägstellung ergebende Reihe:

$$J = -\frac{1}{i\lambda}\gamma + \frac{1}{\lambda^2}\gamma' + \frac{1}{i\lambda^3}\gamma'' - \frac{1}{\lambda^4}\gamma''' \dots$$

für hinreichend große $|u|$ ($|u| \gg 1$) und unter der vereinfachenden Voraussetzung $c' = 0$ aufstellt. Denn für große $|u|$ gilt näherungsweise

$$\frac{d^n \gamma}{d\psi^n} = \gamma^{(n)} = (-1)^n \frac{(1+3)(1+3.2)\dots(1+3n)}{|u|^{4+3n}}.$$

Für hinreichend große λ können also wohl die ersten Glieder die Lösung mit vorgeschriebener Genauigkeit darstellen, aber bei fortschreitender Entwickelung muß das Restglied, ebenso wie das allgemeine Glied $\gamma^{(n)}/\lambda^n$, über alle Grenzen wachsen. Für kleinere Werte von λ dagegen wird im absteigenden Ast schon das erste Restglied nicht mehr vernachlässigt werden können. In dem Fall wird in der Tat, wie die folgende Untersuchung lehrt, außer den bisher aufgestellten Pendelungen konstanter Amplitude noch eine weitere merkliche Pendelung vorhanden sein, die sogar vorwiegend den Charakter der Bewegung bestimmt.

§ 3. Die eingeprägte Pendelung.

Indem wir wieder zur unabhängigen Variabeln u zurückkehren, schreiben wir nach (6′) und (9) die Lösung (12″), in der Form:

$$Q = N e^{i\lambda\psi(u)} \int_{u_0}^{u} \frac{1}{1+u^2} e^{-i\lambda\psi(u)}\, du$$

(12''') $$\psi(u) = \int_0^u \frac{1+u^2}{(1+c'\chi(u))^{\frac{3}{2}}}\, du$$

$$\chi(u) = 2 \cdot \int_0^u \sqrt{1+u^2}\, du = \mathfrak{Ar}\,\mathfrak{Sin}\, u + \sqrt{1+u^2}.$$

Die im vorigen Paragraphen gefundenen Pendelungen waren durch den Anfangswert der Funktion γ, nämlich γ_0, bestimmt, die Schrägstellung durch den Wert $\gamma(u)$. Da $\gamma(u)$ für $u = \pm\infty$ mit seinen Differentialquotienten verschwindet, so könnten wir die erstere auch durch Ausführung der Integration in (12''') von u_0 bis $-\infty$, und die letztere durch Integration von $+\infty$ bis u und Anwendung der Reihenentwickelung des vorigen Paragraphen gewinnen. Es liegt daher nahe, den in (12''') geforderten Integrationsweg von u_0 bis u in 3 Teile zu zerlegen:

$$J_1 = \int_{u_0}^{-\infty}; \quad J_2 = \int_{-\infty}^{+\infty}; \quad J_3 = \int_{+\infty}^{u}.$$

J_1 und J_3 gehen dann in (12''') als Pendelungsamplitude, bezw. als mittlere Schrägstellung ein. J_2 stellt gleichfalls eine Konstante, also wegen des Faktors $e^{i\lambda\psi(u)}$ eine Pendelungsamplitude dar; da sie vom Anfangswert u_0 nicht unmittelbar abhängig ist, sei sie als die „eingeprägte Pendelung“ bezeichnet. Sie bestimmt zunächst die Pendelungsamplitude am Ende der Bahn. Da aber, wie schon früher erwähnt wurde, die Pendelungen im Wesentlichen in der Nähe des Scheitels entstehen werden, dagegen im absteigenden Ast mehr konstant bleiben, so wird sie schon einen guten Anhalt für die eigentlich kritische Amplitude in der Nähe des Scheitels geben.

Die vorangehenden Festsetzungen gelten zunächst nur für den Fall $c' = 0$. Wenn c' statt dessen eine kleine Zahl ist (vgl. S. 5), so sind statt der Grenzen $\pm\infty$ solche Grenzen zu wählen, für die $1+u^2$ groß, aber $1+c'\chi(u)$ noch wenig von 1 verschieden ist. In diesem Sinne sind im Folgenden die Grenzen $\pm\infty$ aufzufassen. Diese Abänderung hat aber keinen Einfluß auf das Resultat, weil sich zeigen wird, daß von dem Integrationsweg nur die Teile in der Nähe des Scheitels wesentlich in Betracht kommen.

Diesen Integrationsweg wollen wir in geeigneter Weise durch komplexes Gebiet führen. Aus der Gleichung (12''') und der

folgenden ergibt sich: Der Exponent in dem Integranden ist für reelle u rein imaginär. Sein reeller Teil fällt zu negativen Werten ab, wenn u in Gebiete mit negativ imaginärem Teil geht (da $d\psi/du$ längs der reellen Axe überall > 0,) und zwar um so rascher, je größer $|u|$. Mit wachsendem imaginärem Teil von u aber wird in $d\psi/du$ das Glied mit u^2 überwiegen, das schließlich negativ werden muß: der reelle Teil des Exponenten wird dann wieder anwachsen. Also: Der absolute Betrag der Exponentialfunktion in (12‴) hat unterhalb der reellen Axe (negativ imaginäre Teile von u) ein Tal, das am tiefsten für große reelle Teile von u, am flachsten in der imaginären Axe ist. Dort liegt also ein S a t t e l p u n k t der Funktion. Er ergibt sich aus $d\psi/du = 0$ zu $u = -i$. In diesem Sattelpunkt liegt aber ein Pol des zweiten Faktors des Integranden, $1/1+u^2$, sodaß der Sattel in 2 Sättel zerlegt wird. Diese ergeben sich, indem wir den Integranden schreiben:

(15) $$e^{\Phi} = e^{-i\lambda\psi(u) - \log(1+u^2)}$$

aus

$$\frac{d\Phi}{du} = -i\lambda\frac{d\psi}{du} - \frac{2u}{1+u^2} = -i\lambda\frac{1+u^2}{[1+c'(\chi(u))]^{\frac{3}{2}}} - \frac{2u}{1+u^2} = 0.$$

Die Wurzeln, die bei hinreichend großem λ, in der Nähe von $-i$ liegen, ergeben sich dann aus folgender Näherungsgleichung, in der wir noch abkürzend

$$\frac{\lambda}{(1+c'\chi(-i))^{\frac{3}{2}}} = \frac{\lambda}{(1-ic'\pi/2)^{\frac{3}{2}}} = \lambda'$$

setzen:

(15′) $$(u^2+1)^2 = \frac{2}{\lambda'}$$

zu

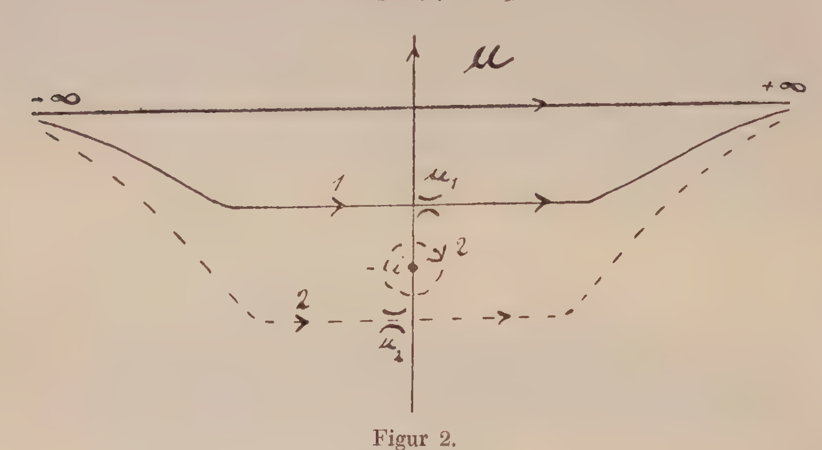
(15″) $$u_1 = -i + \frac{i}{\sqrt{2\lambda'}}, \quad u_2 = -i - \frac{i}{\sqrt{2\lambda'}}$$

Es ist nun naheliegend, den komplexen Integrationsweg in J_2 längs des Tales und über einen dieser Sattelpunkte zu führen (Weg 1 in Figur 2). Das Resultat dieser „Methode der Sattelpunktsintegration“ [1]) ist im allgemeinen das folgende: In der Nähe

1) Vgl. z. B. D e b y e: Näherungsformeln für die Zylinderfunktionen, Math. Ann. 67 (1909), S. 535 und die dort angegebene Literatur.

des Sattels (z. B. u_1) läßt die Funktion Φ sich in eine Potenzreihe entwickeln, deren erste Glieder lauten:

$$\Phi = \Phi_1 + \tfrac{1}{2}\Phi_1''(u-u_1)^2 + \cdots. \tag{16}$$



Figur 2.

Das Integral J_2 läßt sich dann mit einer, mit wachsendem λ wachsenden Genauigkeit ersetzen durch das Integral

$$\int_{-\infty}^{+\infty} e^{\Phi}\,du \approx e^{\Phi_1}\int_{-\infty}^{+\infty} e^{\frac{1}{2}\Phi_1''(u-u_1)^2}\,du = e^{\Phi_1}\sqrt{\frac{2\pi}{-\Phi_1''}}, \tag{16'}$$

falls, (wie hier bereits aus der geometrischen Diskussion ersichtlich ist), Φ_1'' negativ reellen Teil hat. Und zwar stellt dieser Ausdruck im Allgemeinen einen für $\lambda = \infty$ asymptotisch genauen Näherungsausdruck dar.

Im vorliegenden Falle ist diese Schlußweise insofern zu modifizieren, als mit wachsendem λ der Sattelpunkt nach (15″) an den Pol $u = -i$ heranrückt und infolgedessen die Entwickelung (16) aufhört, einen endlichen Gültigkeitsbereich zu haben. Wir werden trotzdem den Näherungsausdruck (16′) verwenden und (um hier die Entwickelung nicht zu unterbrechen, anhangsweise) zeigen, daß er hier eine untere Grenze für den wahren Wert des Integrals J_2 darstellt.

Wir haben für den Sattelpunkt u_1 bei großen Werten von λ (unter Berücksichtigung von $\chi'(-i) = 0$ nach (15′) bezw. (15″)

$$\Phi_1'' = -i\lambda\psi_1'' - 2\frac{1-u_1^2}{(1+u_1^2)^2} = -i\lambda\frac{2u_1}{(1+c'\chi(-i))^{\frac{3}{2}}} - 2\lambda' = -4\lambda'.$$

Wir haben ferner:

$$e^{\Phi_1} = \frac{1}{1+u_1^2} e^{-i\lambda\psi_1} = +\sqrt{\frac{\lambda'}{2}}\, e^{-i\lambda\left[\psi(-i) + \frac{1}{2}(u_1+i)^2\psi''(-i)\right]}$$

Hierin ist nach (15″) und (12‴):

$$-\frac{i\lambda}{2}(u_1+i)^2\psi''(-i) = \frac{i\lambda}{4\lambda'}\,\frac{-2i}{(1-ic'\pi/2)^{-\frac{3}{2}}} = \frac{1}{2}.$$

Um endlich $\psi(-i)$ zu berechnen, beschränken wir uns auf lineare Glieder in c' und erhalten so durch elementare Integration

$$\psi(-i) = \int_0^{-i}(1+u^2)\left(1-\frac{3}{2}c'\chi(u)\right)du$$

$$= \left(u+\frac{u^3}{3}\right)\left(1-\frac{3}{2}c'\chi(u)\right)\Big|_0^{-i} + \frac{3}{2}c'\int_0^{-i}\left(u+\frac{u^3}{3}\right)\sqrt{1+u^2}\,du$$

$$= -\frac{2}{3}i\left(1-ic'\frac{3\pi}{4}\right) - \frac{13}{30}c' = -\frac{2}{3}i - 2c'.$$

Also im Ganzen:

$$e^{\Phi_1} = \sqrt{\frac{\lambda'}{2}}\, e^{-(\frac{2}{3}-2ic')\lambda + \frac{1}{2}}.$$

Schließlich ergibt der Näherungsausdruck (16′) das Resultat:

$$J_2 \approx \frac{\sqrt{\pi}}{2} e^{\frac{1}{2}}\, e^{-(\frac{2}{3}-2ic')\lambda}. \tag{17}$$

Ein analoger Ausdruck wird sich ergeben, wenn man statt des Sattelpunktes u_1 den Sattelpunkt u_2 für den Integrationsweg wählt, (der gestrichelte Weg 2 in Figur 2). Um aber den Integrationsweg über diesen Sattel u_2 legen zu können, muß man ihn vorher über den Pol $u = -i$ hinwegziehen, an dem also ein im Uhrzeigersinn geführtes Residuumintegral bleibt. Dieses letztere ergibt:

$$J_2' = -\operatorname*{Res}_{u=-i}\left\{\frac{1}{1+u^2} e^{-i\lambda\psi(u)}\right\}$$

$$= -\operatorname*{Res}_{u=-i}\left\{\frac{1}{u+i}\right\}\frac{1}{-2i} e^{-i\lambda\psi(-i)}$$

$$= -2\pi i\cdot\frac{i}{2} e^{-(\frac{2}{3}-2ic')\lambda} = \pi e^{-(\frac{2}{3}-2ic')\lambda}.$$

Das Sattelintegral für den Sattel u_2 ist ebenso wie oben das für den Sattel u_1 zu berechnen, wobei sich nur der Unterschied

ergibt, daß jetzt $1/1 + u_2^2 = -\sqrt{\lambda'/2}$ (statt oben $+\sqrt{\lambda'/2}$) zu setzen ist. Man erhält also für dieses Sattelintegral

$$J_2'' \approx -\frac{\sqrt{\pi}}{2} e^{\frac{1}{2}} e^{-(\frac{2}{3} - 2ic')\lambda}$$

und im Ganzen für J_2 den neuen Näherungswert:

$$(17') \qquad J_2 \approx \left(\pi - \frac{\sqrt{\pi}}{2} e^{\frac{1}{2}}\right) e^{-(\frac{2}{3} - 2ic')\lambda}.$$

(Bei der letzteren Berechnung müßte, bei Berücksichtigung der c'-Glieder, darauf Rücksicht genommen werden, daß die Funktion $\chi(u)$ bei $u = -i$ einen Verzweigungspunkt hat. Da aber an der Verzweigungsstelle $\chi'(u) = 0$ und für unsere Berechnung höhere Ableitungen von $\chi(u)$ nicht in Betracht kommen, so hat die Verzweigung keinen Einfluß auf das Resultat.)

Die im Anhang zu beweisende Tatsache, daß die Formel (17) eine untere Grenze für das Integral J_2 darstellt, führt zu dem Schluß, daß der Ausdruck (17′) eine obere Grenze ergibt. Der wahre Wert J_2 ist damit in enge Grenzen eingeschlossen, die numerisch lauten:

$$(17'') \qquad 1{,}47\, e^{-(\frac{2}{3} - 2ic')\lambda} < J_2 < 1{,}67\, e^{-(\frac{2}{3} - 2ic')\lambda}\,[1].$$

Es ist bemerkenswert, daß der absolute Beitrag dieser Amplitude, bei unserer Definition von λ, von c' unabhängig, also der gleiche wird, als ob man als ballistische Kurve eine widerstandsfreie Wurfparabel mit derselben Scheitelgeschwindigkeit zugrunde legte. Nur die Phase der durch J_2 bestimmten Pendelung ist von c' abhängig.

Um ein numerisches Beispiel zu geben, sei angenommen: Abgangswinkel $\varphi_0 = 60^0$, $\lambda = 4$, $c' = 0$. Dann wird die mittlere Schrägstellung γ/λ im Maximum für $u = 0$ (Scheitel) zu $1/\lambda = 0{,}25 \approx 14^0$; die durch die Erhöhung bestimmte Pendelungsamplitude $J_1 = \gamma_0/\lambda = 1/(1 + u_0^2)^2 \lambda = 1/64 \approx 1^0$; dagegen die „eingeprägte Pendelungsamplitude“ J_2, als Mittelwert aus den Grenzen (17″) berechnet $= \pi/2\ e^{-\frac{2}{3}\lambda} = 0{,}11 \approx 6{,}5^0$, sie überwiegt somit wesentlich die Amplitude J_1.

Der gesamte Ausschlag kann in diesem Falle maximal $14^0 + 6{,}5^0 + 1^0 = 21{,}5^0$ betragen, sodaß an der Zulässigkeit unserer

---

1) Genauer gesagt: Diese Ungleichungen gelten für den Betrag der Amplitude J_2, während die Phase durch den Faktor $e^{2ic'\lambda}$ dargestellt ist.

Voraussetzung II nicht zu zweifeln ist. Für weitere numerische Angaben, insbesondere über die Abhängigkeit der Größe λ von den Bahnelementen (Anfangsgeschwindigkeit und Erhöhung, sowie Geschoßform und Kaliber), sei auf die gleichnamige Veröffentlichung des Verfassers in den Artilleristischen Monatsheften, Mai-Juni-Heft 1919, verwiesen.

Wir haben noch den Beweis nachzutragen, daß der Näherungsausdruck (17) eine untere (und entsprechend (17′) eine obere) Grenze der Amplitude J_2 darstellt. Zu dem Zweck haben wir den Fehler dieser Darstellung zu berechnen. Mit $u - u_1 = w$ wird der genaue Wert von J_2:

$$= \int_{-\infty}^{+\infty} \frac{1}{1+u^2} e^{-i\lambda\psi(u)}\, du = \int_{-\infty}^{+\infty} \frac{e^{-i\lambda[\psi_1 + \psi_1' w + \frac{1}{2}\psi_1'' w^2 + \frac{1}{6}\psi_1''' w^3 \ldots]}}{(u_1 + i + w)(u_1 - i + w)}\, dw$$

$$= \frac{e^{-i\lambda\psi_1}}{1+u_1^2} \int_{-\infty}^{+\infty} \frac{e^{-\sqrt{2\lambda'}\,w - \lambda' w^2 - i\lambda'(w^3/3 + \cdots)}}{(1 - i\sqrt{2\lambda'}\,w)(1 + iw/2)}\, dw.$$

Wir setzen: $\sqrt{\lambda'}\,w = y$ und unterdrücken, für große λ', alle negativen Potenzen von λ' neben endlichen Gliedern. Es wird:

$$J_2 = \frac{e^{\Phi_1}}{\sqrt{\lambda'}} \int_{-\infty}^{+\infty} \frac{e^{-i\sqrt{2}\,y - y^2}}{1 - i\sqrt{2}\,y}\, dy.$$

Andererseits ist nach (16′) unser Näherungsausdruck (17) für J_2:

$$e^{\Phi_1} \int_{-\infty}^{+\infty} e^{\frac{1}{2}\Phi_1'' w^2}\, dw = \frac{e^{\Phi_1}}{\sqrt{\lambda'}} \int_{-\infty}^{+\infty} e^{-2y^2}\, dy.$$

Der relative Fehler dieses Ausdrucks ist in der Tat endlich, nämlich:

$$\text{(18)}\quad \begin{aligned} \frac{\Delta J_2}{J_2} &= \frac{1}{J_2}\frac{e^{\Phi_1}}{\sqrt{\lambda'}} \int_{-\infty}^{+\infty} e^{-y^2}\left(\frac{e^{-i\sqrt{2}y}}{1 - i\sqrt{2}\,y} - e^{-y^2}\right) dy \\ &= \frac{2\sqrt{2}}{\sqrt{\pi}} \int_0^{\infty} e^{-y^2}\left(\frac{\cos\sqrt{2}\,y + \sqrt{2}\,y \sin\sqrt{2}\,y}{1+2y^2} - e^{-y^2}\right) dy. \end{aligned}$$

Die Klammer in diesem Integral verschwindet quadratisch für $y = 0$ und hat oszillierendes Vorzeichen, und zwar liegen die Vorzeichenwechsel annäherd an den Punkten $\sqrt{2}\,y = n\pi$. Wegen

des rasch abnehmenden Faktors e^{-y^2} ist das Vorzeichen der Klammer im ersten Intervall (0 bis π) maßgebend für das Vorzeichen des Fehlers ΔJ_2 und dieses ist positiv, wie man durch Berechnung der höheren (als zweiten) Ableitungen nach y bei $y = 0$ oder Berechnung in einzelnen Punkten (z. B. $\sqrt{2}y = \pi/2$) leicht sieht. Der Näherungsausdruck (17) ist also eine untere Grenze (bei großem λ); da in (17′) der nämliche Ausdruck mit negativem Vorzeichen auftritt, ist (17′) eine obere Grenze für den wahren Wert J_2. Die Fehler können übrigens nach (18) leicht numerisch berechnet werden.

Die in der Einleitung gestellten Fragen seien noch kurz beantwortet. Die gefragte Größenverbindung, von der die Präzessionspendelungen abhängen, ist die Größe λ (Gleichung 8). Bezeichnet u_0, u_e Anfangs- und Endwert der Variabeln u, so ist nach (12′):

$$n = \lambda \frac{\psi(u_e) - \psi(u_0)}{2\pi}$$

die Anzahl der Pendelungen längs der Bahn. Ihre Größe ist im Wesentlichen durch die kleine Amplitude $|J_1| = \gamma_0/\lambda$ im aufsteigenden Ast der Bahn, und durch $|J_2|$ (17″) im absteigenden Ast gegeben. Die letztere ist für große λ unwesentlich, sie wächst aber, wenn λ abnimmt und erreicht Werte, für die unsere Voraussetzung II sicher nicht mehr zulässig ist, wenn etwa $\lambda \leqq 2$. Durch qualitative Überlegungen kann man zeigen, daß für so kleine λ die Gefahr der Nichtfolgsamkeit des Geschosses und des Bodentreffers besteht. Neben die Stabilitätsforderung (s. Voraussetzung III) tritt somit als *Bedingung der Folgsamkeit, daß λ groß, mindestens > 2 sein muß.*

Die ballistische Kurve ist auf Grund des Einflusses der Schrägstellung und der Pendelungen zu verbessern. Für den Widerstand gilt dabei weiter das quadratische Gesetz, aber mit einem, insbesondere im Scheitel und im absteigenden Ast, vergrößerten Widerstandskoeffizienten, der aus dem, aus Schrägstellung und Pendelung berechneten mittleren Anstellwinkel auf Grund empirischer Gesetze zu bestimmen ist. Schließlich kann die Seitenabweichung leicht berechnet werden, wenn die (veränderliche) Pendelung J_2 klein ist neben der mittleren Schrägstellung γ/λ (für $\lambda > 5$ ist $J_2 < \frac{1}{3}\gamma_{\max}/\lambda$). Für die ablenkende Kraft (den „Auftrieb“) an dem bewegten Geschoß gilt nämlich, genügend kleine ϑ vorausgesetzt, ein analoges Gesetz wie für das Widerstandsmoment:

$$A = kv^2 \sin\vartheta,$$

worin k empirisch bestimmbar ist. Für die horizontale Seitenablenkung Z gilt dann mittels (13′) und der Definitionsgleichungen von γ und λ:

$$\frac{d^2 z}{dt^2} = kv^2 \frac{Z}{N} = kv^2 \frac{\gamma}{\lambda} = \frac{k}{C} Ng \frac{\cos\varphi}{v}. \tag{19}$$

Durch Einführung der unabhängigen Variabeln u wird diese Gleichung integrabel.

Weitere Ausführungen dieser Folgerungen ersehe man in der oben zitierten Arbeit in den Artilleristischen Monatsheften. Die Gleichung (19) stimmt mit der von Cranz (Lehrb. Bd. I S. 377) auf mehr qualitativ geometrischem Wege aufgestellten überein. Vermutlich wird eine ähnliche Gleichung auch für größere Geschwindigkeiten (also allgemeinere Widerstandsgesetze) bestehen, doch scheint mir die unmittelbare Übertragung der Gleichung (19), die bei Cranz gemacht wird, zweifelhaft, da sie voraussetzt, daß der Auftrieb A (und damit auch das Moment M) in gleicher Weise von v abhängen, wie der Widerstand W, worüber weder empirisch noch theoretisch etwas bekannt ist.

Über die Zerlegung total positiver Zahlen in Quadrate.

Von

Edmund Landau.

Vorgelegt in der Sitzung vom 4. Juli 1919.

Herr Hilbert[1]) hat vor 20 Jahren ohne Ausführung seines Beweises den Satz ausgesprochen: *Jede total positive Zahl eines algebraischen Zahlkörpers*[2]) *läßt sich als Summe von vier Quadraten gewisser Zahlen des Körpers darstellen.* Ein Beweis dieser Behauptung ist bisher weder von ihm noch von anderer Seite veröffentlicht worden, und ich kann nicht feststellen, ob der Satz richtig ist. Aber ich kann jetzt drei kleine Beiträge liefern. Ich werde

I) den Hilbertschen Satz für jeden quadratischen Körper beweisen;

II) für jeden Körper beweisen, daß jede total positive Zahl überhaupt als Summe von Quadraten (von Körperzahlen) darstellbar ist;

III) mit Hilfe von II) eine scheinbar einfachere Formulierung (die nur von rationalzahligen Polynomen spricht) angeben, mit deren Richtigkeit der Hilbertsche Satz steht und fällt.

I) Es sei ein quadratischer Körper $P(\sqrt{D})$ der Grundzahl D gegeben, $\alpha = A + B\sqrt{D}$ (A, B rational) eine total positive Zahl des Körpers, ohne Beschränkung der Allgemeinheit[3]) $\alpha \neq 0$; α'

---

1) *Grundlagen der Geometrie*; § 38 in allen Auflagen.

2) Sind alle konjugierten Körper imaginär, so heißt jede Zahl α des Körpers total positiv. Anderenfalls heißt α total positiv, wenn die zu α konjugierten Zahlen in den reellen konjugierten Körpern positiv sind.

3) $0 = 0^2 + 0^2 + 0^2 + 0^2$; übrigens nennt Herr Hilbert 0 nur im Falle $D < 0$ total positiv.

bezeichne die zu α konjugierte Zahl in $P(-\sqrt{D})$. Dann sind $S(\alpha) = \alpha + \alpha'$ (Spur) und $N(\alpha) = \alpha\alpha'$ (Norm) rational, speziell $N(\alpha) > 0$ (da $\alpha \neq 0$ und im Falle $D > 0$ sowohl α als α' positiv ist). Nach Lagrange ist $N(\alpha) = a^2 + b^2 + c^2 + d^2$, wo a, b, c, d rational sind und ohne Beschränkung der Allgemeinheit $a^2 > 0$, also (da a durch $-a$ ersetzt werden kann) $a \neq A$ gewählt werden kann. Dann ergibt sich

$$2\alpha(A-a) = \alpha(S(\alpha)-2a) = \alpha(\alpha+\alpha'-2a) = \alpha^2 + N(\alpha) - 2\alpha a$$
$$= \alpha^2 + a^2 + b^2 + c^2 + d^2 - 2\alpha a = (\alpha - a)^2 + b^2 + c^2 + d^2 = \sum_{\nu=1}^{4} \beta_\nu^2,$$

wo die β_ν Zahlen des Körpers sind, also unter Benutzung der bekannten Identität über das Produkt zweier Summen von je 4 Quadraten

$$\alpha \operatorname{sign}(A-a) = \frac{1}{2\,|A-a|} \sum_{\nu=1}^{4} \beta_\nu^2 = \sum_{\nu=1}^{4} b_\nu^2 \sum_{\nu=1}^{4} \beta_\nu^2 = \sum_{\nu=1}^{4} \gamma_\nu^2.$$

Im Falle $A > a$ ist damit die Behauptung bewiesen; im Falle $A < a$ ist $D < 0$ (da sonst $D > 0$, $\alpha \leqq 0$ wäre), also

$$\alpha = -\sum_{\nu=1}^{4} \gamma_\nu^2 = |D| \left(\frac{\sqrt{D}}{|D|}\right)^2 \sum_{\nu=1}^{4} \gamma_\nu^2 = |D| \sum_{\nu=1}^{4} \delta_\nu^2$$
$$= \sum_{\nu=1}^{4} c_\nu^2 \sum_{\nu=1}^{4} \delta_\nu^2 = \sum_{\nu=1}^{4} \varepsilon_\nu^2.$$

II) Es sei ein Körper $n$ten Grades $P(\vartheta)$ gegeben, ϑ Wurzel der irreduziblen Gleichung

$$f(x) = x^n + a_{n-1} x^{n-1} + \cdots + a_0 = 0$$

mit rationalen Koeffizienten, $\alpha \neq 0$ eine total positive Zahl von $P(\vartheta)$. Dann ist $\alpha = g(\vartheta)$, wo $g(x)$ ein rationalzahliges Polynom von kleinerem als $n$tem Grade ist. Es genügt, die Existenz zweier definiter [4]) rationalzahliger Polynome $h_1(x)$ und $h_2(x)$ anzugeben, sodaß $h_1(x) \not\equiv 0$ (mod. $f(x)$) und

(1) $$h_1(x)\, g(x) \equiv h_2(x) \pmod{f(x)}$$

ist. Denn nach einem Satz von mir [5]) ist

$$h_1(x)\, h_2(x) = \sum_{\nu=1}^{q} k_\nu^2(x),$$

---

4) Damit ist gemeint, daß die Funktion $\geqq 0$ für alle reellen x ist.

5) *Über die Darstellung definiter binärer Formen durch Quadrate* [Mathematische Annalen, Bd. LVII (1903), S. 53—64], S. 61.

wo die k_ν rationalzahlige Polynome sind, also wegen $h_1(\vartheta) \neq 0$

$$h_1(\vartheta)\alpha = h_1(\vartheta)g(\vartheta) = h_2(\vartheta),$$

$$\alpha = \frac{h_1(\vartheta)h_2(\vartheta)}{h_1^2(\vartheta)} = \frac{\sum\limits_{\nu=1}^{q} k_\nu^2(\vartheta)}{h_1^2(\vartheta)} = \sum_{\nu=1}^{q} \eta_\nu^2.$$

(1) beweise ich folgendermaßen. Es seien m unter den zu ϑ konjugierten Zahlen $\vartheta_1, \ldots, \vartheta_n$ reell und im Falle $m > 0$ die Numerierung so gewählt, daß $\vartheta_1, \ldots, \vartheta_m$ reell sind.

1) Im Falle $m = 0$ ist $f(x)$ (für reelle x) stets > 0, also nach dem Hilfssatz in § 1 meiner erwähnten Abhandlung bei passend klein gewähltem rationalem $d > 0$ die Funktion

$$f(x) + d\,g(x) = h_2(x)$$

definit, also (1) mit $h_1(x) = d$ erfüllt.

2) Im Falle $m = n$ setze ich

$$h_2(x) = \sum_{\nu=1}^{n} \frac{g(\vartheta_\nu)}{f'^2(\vartheta_\nu)} \left(\frac{f(x)}{x-\vartheta_\nu}\right)^2.$$

Dies ist ein rationalzahliges Polynom und zwar definit wegen $g(\vartheta_\nu) > 0$ ($\nu = 1, \ldots, n$). Ferner ist für $\nu = 1, \ldots, n$

$$h_2(\vartheta_\nu) = g(\vartheta_\nu),$$

also

$$g(x) \equiv h_2(x) \pmod{f(x)},$$

somit (1) mit $h_1(x) = 1$ erfüllt.

3) Im Falle $0 < m < n$ setze ich

$$\varphi(x) = \prod_{\nu=1}^{m} (x-\vartheta_\nu), \qquad \psi(x) = \prod_{\nu=m+1}^{n} (x-\vartheta_\nu);$$

das sind reellzahlige Polynome, und zwar ist stets $\psi(x) > 0$, sowie

$$\varphi(x)\psi(x) = f(x).$$

Ich setze ferner

$$\chi(x) = (1 + x^{2n}\varphi^2(x)) \sum_{\nu=1}^{m} \frac{g(\vartheta_\nu)}{\varphi'^2(\vartheta_\nu)} \left(\frac{\varphi(x)}{x-\vartheta_\nu}\right)^2.$$

Das ist ein reellzahliges Polynom und zwar stets > 0 wegen $g(\vartheta_\nu) > 0$ ($\nu = 1, \ldots, m$). Ferner ist für $\nu = 1, \ldots, m$

$$\chi(\vartheta_\nu) = g(\vartheta_\nu),$$

also

$$\chi(x) = g(x) + \mu(x)\varphi(x),$$

wo $\mu(x)$ ein reellzahliges Polynom ist; folglich

$$\omega(x) = \chi(x)\psi(x) = g(x)\psi(x) + \mu(x)f(x).$$

$\omega(x)$ ist stets > 0, da dies von $\psi(x)$ und $\chi(x)$ galt. Der genaue Grad von $\omega(x)$ ist $2n + 2m + 2(m-1) + (n-m) > n + (n-m)$, also größer als der von $g(x)\psi(x)$; er ist also gleich dem von $\mu(x)f(x)$. Ich ersetze nun nach dem schon oben benutzten Hilfssatz $\psi(x)$ und $\mu(x)$ durch rationalzahlige Polynome $h_1(x)$ und $h_3(x)$ mit bezw. so benachbarten Koeffizienten, daß erstens $h_1(x)$ genau den Grad $n-m$ von $\psi(x)$ behält (also nicht durch $f(x)$ teilbar ist) und definit bleibt, zweitens das aus $\omega(x)$ entstehende rationalzahlige Polynom

$$h_2(x) = g(x)h_1(x) + h_3(x)f(x)$$

auch wieder definit ist. Dann ist (1) erfüllt.

III) Den Hilbertschen Satz nenne ich **H**; unter **L** verstehe ich folgende Aussage (von der ich auch nicht weiß, ob sie richtig oder falsch ist): *$f(x)$ sei ein irreduzibles rationalzahliges Polynom; $a_1(x), \ldots, a_5(x)$ seien fünf rationalzahlige Polynome. Dann gibt es vier rationalzahlige Polynome $b_1(x), \ldots, b_4(x)$ derart, daß*

$$\sum_{\nu=1}^{5} a_\nu^2(x) \equiv \sum_{\nu=1}^{4} b_\nu^2(x) \pmod{f(x)} \tag{2}$$

ist.

1) *Wenn* **L** *richtig ist, so ist* **H** *richtig.* Denn α sei eine total positive Zahl des Körpers $P(\vartheta)$, $f(x) = x^n + \cdots + a_0$ das zu ϑ gehörige irreduzible Polynom. Nach II) ist:

$$\alpha = g(\vartheta) = \sum_{\nu=1}^{q} \eta_\nu^2,$$

also bei passender Wahl rationalzahliger Polynome $e_\nu(x)$

$$g(x) \equiv \sum_{\nu=1}^{q} e_\nu^2(x) \pmod{f(x)}.$$

Ist $q > 4$, so ist nach **L**

$$\sum_{\nu=1}^{5} e_\nu^2(x) \equiv \sum_{\nu=1}^{4} l_\nu^2(x) \pmod{f(x)},$$

also $g(x)$ einer Summe von $q-1$ Quadraten kongruent, u. s. f. bis zu der (im Falle $q \leq 4$ von vornherein gültigen) Kongruenz

$$g(x) \equiv \sum_{\nu=1}^{4} f_\nu^2(x) \pmod{f(x)};$$

aus ihr folgt

$$\alpha = g(\vartheta) = \sum_{\nu=1}^{4} f_\nu^2(\vartheta) = \sum_{\nu=1}^{4} \lambda_\nu^2.$$

2) *Wenn* **H** *richtig ist, so ist* **L** *richtig.* Denn unter den Voraussetzungen von **L** ist, wenn ϑ eine Wurzel von $f(x) = 0$ ist, $\alpha = \sum_{\nu=1}^{5} a_\nu^2(\vartheta)$ eine total positive Zahl in $P(\vartheta)$ oder $= 0$, also nach **H**

$$\alpha = \sum_{\nu=1}^{4} \lambda_\nu^2 = \sum_{\nu=1}^{4} b_\nu^2(\vartheta),$$

folglich (2) erfüllt.

Leitfähigkeit und Jonengehalt der Atmosphäre im geschlossenen Raum bei konstanten Licht- und Temperaturverhältnissen.

Von

Dr. Rose Stoppel.

Vorgelegt von E. Wiechert in der Sitzung am 28. März 1919.

Die vorliegenden Beobachtungen über Leitfähigkeit und Jonengehalt der Atmosphäre wurden angeregt durch meine Untersuchungen über periodische Erscheinungen bei Pflanzen[1]. Es hatte sich gezeigt, daß der sogenannte Pflanzenschlaf — wenigstens bei den beobachteten Laubblättern — außer von dem Wechsel von Licht- und Temperaturintensitäten abhängig ist von einem tagesperiodisch sich ändernden Faktor, der elektrischer Natur sein muß.

Es ist bekannt, daß die Potentialdifferenz und die Leitfähigkeit der Atmosphäre tagesperiodischen Schwankungen unterliegt. Die Potentialdifferenzen kamen aber bei den pflanzlichen Versuchen als Erreger deshalb nicht in Betracht, weil die Versuchspflanzen an elektrisch geschützten Orten, im Keller, Laboratorium oder im Gewächshaus standen. Die Periodizität der Leitfähigkeit dagegen, die in der freien Atmosphäre an verschiedensten Orten beobachtet worden war, hat eine Kurve, die eine große Ähnlichkeit zeigt mit der, die sich aus der wechselnden Stellung der Blätter bei den Schlafbewegungen ergab. Es lag daher nahe, in den Intensitätsschwankungen der Leitfähigkeit den mutmaßlichen Faktor zu sehen, der den Schlaf der Blätter reguliert. Einen Beweis für diese Annahme mußte ich zunächst aus Mangel an den technischen Hilfs-

---

1) R. Stoppel, Die Abhängigkeit der Schlafbewegungen von Phaseolus multiflorus von verschiedenen Außenfaktoren. Zeitschr. f. Bot. 8, 1916, S. 609—684.

mitteln schuldig bleiben. Die Vermutung stand auch darum noch auf schwankenden Füßen, weil die Periodizität der Leitfähigkeit in geschlossenen Räumen bei im übrigen möglichst konstanten Außenverhältnissen, wie es bei meinen Pflanzenversuchen der Fall gewesen, noch nicht nachgewiesen worden war. Im Gegenteil, wenn der Jonengehalt der Atmosphäre in hohem Maße abhängig wäre von dem Luftdruck und seinen Schwankungen, wie es die Ebertsche Theorie der Bodenatmung verlangt, so war es anzunehmen, daß die Kurve der Leitfähigkeit in geschlossenen Räumen keine feste Rhythmik aufweisen würde besonders in Hamburg, wo die Barometerkurve außerordentlich unregelmäßig ist.

Es war daher mein Wunsch, diese Lücke in unseren Kenntnissen über die Schwankungen der Leitfähigkeit und des Jonengehaltes der Atmosphäre in geschlossenen Räumen auszufüllen. Herr Professor Classen in Hamburg hatte die große Liebenswürdigkeit, mir die erforderlichen Apparate aus dem physikalischen Institut in entgegenkommendster Weise für die Untersuchungen zur Verfügung zu stellen. Die Arbeitsmöglichkeit und den Arbeitsraum erhielt ich durch Herrn Professor Winkler. Beiden Herren spreche ich für ihre Bereitwilligkeit und das der Arbeit entgegengebrachte dauernde Interesse meinen wärmsten Dank aus.

Der Arbeitsraum, ein Kellerraum des Instituts für allgemeine Botanik, diente ausschließlich den vorliegenden Versuchen und wurde während der Versuchszeiten nur von mir betreten. Er liegt unter der Treppe des Institutseingangs und steht nicht direkt mit der Außenluft in Verbindung. Er hat drei Türen. Zwei derselben führen zu Nebengelassen, deren Fenster stets geschlossen blieben. Diese Zwischentüren wurden während der Versuchszeiten ebenfalls nicht geöffnet. Die dritte Tür, durch die ich den Versuchsraum betrat, verband denselben mit dem langen Mittelgang des Kellers. In dem Versuchsraum selbst befand sich weder eine Gas- noch Wasserleitung, auch keine Heizrohre. Nur eine kurze isolierte elektrische Leitung führte zu einer Deckenbeleuchtung. An diese schloß ich vermittelst Schraubkontakt den Leitungsdraht an, der zu den später zu besprechenden elektrischen Lampen führte. Die Lampen standen auf einem niedrigen Bord, während die Apparate auf einem Holztisch Aufstellung fanden. Ich achtete darauf, daß der Leitungsdraht der Lampen nicht mit dem Tisch in Berührung kam. Der Fußboden des Raumes ist betoniert, darüber ist eine Zwischenschicht aus Holzfaser, obenauf Linoleum. Die Decke ist von eisernen Trägern durchzogen, auf denen die Treppenstufen des Instituteingangs ruhen. Decke und Wände

sind mit Kalk abgeputzt. Gegen Erschütterung und Staub ist der Raum gut geschützt.

Ich achtete darauf, daß meine Garderobe während der Beobachtungszeiten möglichst die gleiche blieb, um die durch Wolle und Seide entstehenden Versuchsfehler zu vermeiden.

An Instrumenten stand mir für die Untersuchung der Leitfähigkeit ein Exner'sches Aluminiumblatt-Elektroskop nach Elster und Geitel mit Lupenablesung zur Verfügung. Da die Zerstreuung in dem Kellerraum groß war, mußte ich mich mit einem 164 mm langen, 4 mm dicken, geschwärzten Metallstab als Zerstreuungskörper begnügen. Trotzdem konnte ich nur eine Beobachtungszeit von 5 Minuten ansetzen, und mußte sogar die erste Ablesung schon wenige Sekunden nach dem Aufladen vornehmen, sobald die Blättchen in Ruhe gekommen waren. Nur die Novemberbeobachtungen hatte ich mit einer Beobachtungszeit von 10 Minuten durchgeführt. Die in der Kurve eingetragenen Werte dieses Monats sind durch Division aus den gefundenen erhalten. In der Folge blieb ich für diesen Apparat bei einer Versuchsdauer von 5 Minuten.

Nach erfolgter Ablesung verließ ich möglichst ruhig und schnell den Versuchsraum und betrat ihn erst wieder nach Ablauf der Versuchszeit, die mit Hilfe einer Stoppuhr kontrolliert wurde. Bei den Ablesungen notierte ich die Stellung beider Aluminiumblättchen und achtete darauf, daß während des Ablesens mein Kopf dem Zerstreuungskörper nicht zu nahe kam. Einen Einfluß der Haare eliminierte ich dadurch, daß ich die Hand zwischen Kopf und Instrument hielt. Da ich alle Handgriffe jedesmal pedantisch gleichartig auszuführen mich bemühte, so sind die durch die direkte Ablesung entstandenen Fehler für die relativen Werte weniger von Bedeutung. Ein Schutzzylinder wurde nicht benutzt, da das Instrument sich an einem elektrisch völlig geschützten Ort befand, und die gefundenen Werte ein brauchbares Maß für die Größe der Leitfähigkeit der Atmosphäre abgeben sollten.

Der Jonenaspirator, der auf dem gleichen Tisch stand, war von dem freistehenden Elektroskop durch ein aufrecht gestelltes Brett getrennt, damit der durch den Aspirator erzeugte Luftstrom nicht direkt auf den Zerstreuungskörper wirkte. Der Jonenaspirator stammt ebenfalls aus der Werkstatt von Günther und Tegetmeyer, Braunschweig und diente zur Bestimmung der Jonenmenge. Leider ist der Apparat noch nicht mit einem Anemometer versehen, und die Fördermenge des Uhrwerks ist mir nicht bekannt. Zudem weiß ich nicht, ob dieselbe während der ganzen

Versuchszeit konstant blieb, da das Uhrwerk zweimal reparaturbedürftig war. Im Dezember fielen deshalb die Untersuchungen ganz aus, die zweite Reparatur einer Feder erfolgte nach Beendigung der Ablesungen im März. Absolute Berechnungen über den Jonengehalt der Atmosphäre konnte ich deshalb nicht anstellen, doch dürften die gefundenen Werte dennoch ein brauchbares relatives Maß abgeben. — Die Länge der inneren Elektrode betrug 40 cm, der Durchmesser der äußeren 3 cm. Die Beobachtungszeit erstreckte sich auf 10 Minuten. Das Instrument war durch Anschluß an eine Wasserleitung geerdet.

Bei beiden Instrumenten wurden nur Untersuchungen mit + Ladung der Elektroden gemacht, also die — Leitfähigkeit resp. der — Jonengehalt bestimmt. Wiederholte Prüfungen ergaben, daß die Isolation der Apparate gut war. Herr Professor Classen ließ mir die Instrumente im elektrischen Prüfungsamt durch Herrn Buchheim aichen, wofür ich auch an dieser Stelle nochmals danke.

Zum Aufladen der Elektroskope diente eine Zambonisäule Ich bemühte mich, die Anfangspotentiale immer möglichst auf die gleiche Höhe zu bringen, was jedoch nur in beschränktem Maße gelang, da ich meinen Aufenthalt in dem Versuchsraum nicht zu sehr ausdehnen wollte. So schwankten die Anfangsladungen beider Instrumente im allgemeinen zwischen 200 und 220 Volt.

Die Temperatur wurde durch direkte Ablesung an einem Thermometer mit $^{1}/_{10}{}^{0}$ Einteilung bestimmt. Mehrfach stellte ich ein selbstregistrierendes Psychrometer in dem Versuchsraum auf. Da diese Kurven innerhalb jeder Versuchsperiode fast gar keine Schwankungen anzeigten, sind die Resultate nicht weiter verwertet.

Ein Barograph zeichnete seine Kurve in einem Zimmer des darüber liegenden Stockwerkes auf. Die Kurven desselben sind mit B bezeichnet in die Kurven der Leitfähigkeitsmessungen eingezeichnet. Leider fehlen im November und April gerade die Angaben von den Versuchstagen, und im Juni ist eine mehrstündliche Unterbrechung der Aufzeichnungen gewesen.

Die Untersuchungen wurden in jedem Monat einige Tage in dauernder Dunkelheit und mehrere Tage bei konstantem Licht angestellt. In der Dunkelperiode wurden die Ablesungen mit Hilfe einer elektrischen Taschenlampe ausgeführt, bei den Lichtversuchen brannten in dem Raum zwei Osramlampen von je 200 Kerzen Stärke. Die Lampen standen auf dem erwähnten Bord neben einander und ebleuchteten beide den Jonenaspirator von der hinteren Längsseite

in ca. 3/4 m Abstand vom Apparat. Das Elektroskop mit freiem Zerstreuungskörper war durch das oben erwähnte Brett so geschützt, daß das ganze Instrument von keinem direkten Lichtstrahl getroffen wurde.

Auch der Gang der durchdringenden Strahlung wurde in dem gleichen Versuchsraum einige Monate hindurch beobachtet. Herr Professor Jensen hatte die Liebenswürdigkeit, mir den diesbezüglichen Apparat so lange zu überlassen. Die Ablesungen erfolgten alle 2 Stunden gleichzeitig mit den übrigen Beobachtungen. Nach jeder Ablesung wurde das Elektroskop neu aufgeladen und wieder 2 Stunden exponiert. Die Resultate dieser Beobachtungen veröffentliche ich nicht, da ich dieselben für nicht genau genug halte, um theoretische Schlüsse auf dieselben aufbauen zu können. So viel kann aber gesagt werden, daß wenn überhaupt eine Periodizität der durchdringenden Strahlung existierte, so kann es sich nur um ein Minimum in den frühen Morgenstunden handeln. Diese Annahme bekommt darum den Wert einer Wahrscheinlichkeit, weil dieselbe Beobachtung auch von anderer Seite gemacht worden ist. Auf alle Fälle sind die Schwankungen auch außerordentlich gering.

Versuchsergebnisse.

Aus praktischen Gründen muß ich von einer Wiedergabe meiner Versuchsprotokolle absehen, doch stehen sie Jedem gerne zur Verfügung, der sich für dieselben interessiert. In den beigegebenen Kurven sind die Versuchsresultate jedoch schon ziemlich vollständig zu übersehen.

Die Kurven sind so zu verstehen, daß auf der Abscisse die Zeiten abgetragen sind und zwar derart, daß die Strecke zwischen 2 Strichen einem Zeitraum von 2 Stunden entspricht. 12 Uhr mittags ist durch die verstärkte Linie kenntlich sowie durch die über jeder Kurve eingetragenen Daten. Auf der Ordinate ist der Ladungsverlust der Instrumente in Volt aufgetragen, der jeweils während der Versuchsdauer eintrat, also bei dem Elektroskop mit freistehendem Zerstreuungskörper innerhalb 5 Minuten, beim Jonenaspirator im Lauf von 10 Minuten. Das Intervall bei der Ordinate zwischen 2 verstärkten Strichen entspricht einer Differenz von 10 Volt. — Die Resultate der Untersuchungen aus den Dunkelperioden sind mit ausgezogenen Linien, die der Lichtperiode mit unterbrochenen wiedergegeben. Die Ablesungen fanden meistens alle 2 Stunden statt. — Die Kurven des Barographen sind mit

B bezeichnet in die Leitfähigkeitskurven eingezeichnet, zu ihnen gehören die Millimeterwerte am Rande rechts. — Jede Kurve ist das Versuchsergebnis der Tage eines Monats.

Die Temperaturkurven sind nicht eingezeichnet, um das Bild nicht verworren zu machen. Da die Temperatur im Versuchsraum sich aber sehr konstant hielt, so genügt die zusammenfassende Tabelle (s. nächste Seite), um sich ein Bild von dem Einfluß der Temperaturschwankungen machen zu können. In der Tabelle sind die höchsten und tiefsten Temperaturen eingetragen, die während jeder Licht- und Dunkelperiode erreicht wurden, außerdem die größte Temperaturschwankung jeder einzelnen Periode innerhalb 24 Stunden von Mitternacht an gerechnet. Unberücksichtigt blieb bei dieser letzten Angabe die Schwankung am ersten 1/2 Tag nach dem Wechsel von Licht- und Dunkelperiode, da sich dann die Temperatur erst allmählich auf ihren neuen Mittelwert einstellte. Um die etwaige Nachwirkung einer vorhergegangenen Periode erkennen zu können, richtete ich es so ein, daß in den ersten 4 Monaten zuerst die Dunkeluntersuchungen einsetzten, in den letzten 4 Monaten die Lichtversuche.

Das Kurvenblatt I S. 404 und III S. 410 gibt die Zerstreuung der Elektrizität an bei freistehendem Zerstreuungskörper, das Blatt II S. 405 und IV S. 411 den Jonengehalt, gemessen mit dem Jonenaspirator.

Bei Besprechung der Versuchsergebnisse wende ich mich zuerst den mit freistehendem Zerstreuungskörper gewonnenen zu. Bei Durchsicht der Kurven fällt zunächst eine deutliche tagesperiodische Rhythmik während der Dunkelperiode auf. Das Maximum der Zerstreuung liegt jeweils in den ersten Morgenstunden und ist meist zwischen 2 und 4 Uhr zu suchen. Genau ist der Wendepunkt der Kurve nicht zu bestimmen, da die Schwankungen der Leitfähigkeit außerordentlich schnell eintreten und oft recht groß sind. Es ist also sehr wohl möglich und sogar wahrscheinlich, daß die eingetragenen Maxima garnicht den tatsächlichen entsprechen, sondern daß diese in den unbeachteten Zwischenstunden erreicht worden sind. Es unterliegt aber keinem Zweifel, daß diese Maximalwerte in der Regel in den frühen Morgenstunden liegen und hierin würden sich dann meine Erfahrungen decken mit den von Kähler[1]) in Potsdam gewonnenen, von Dorno[2]) in Davos,

1) K. Kähler, Die Schwankungen des luftelektrischen Leitvermögens und des vertikalen Leitungsstromes am Erdboden. Ergeb. d. meteorol. Beobachtungen in Potsdam 1911, S. XVI—XXXI.

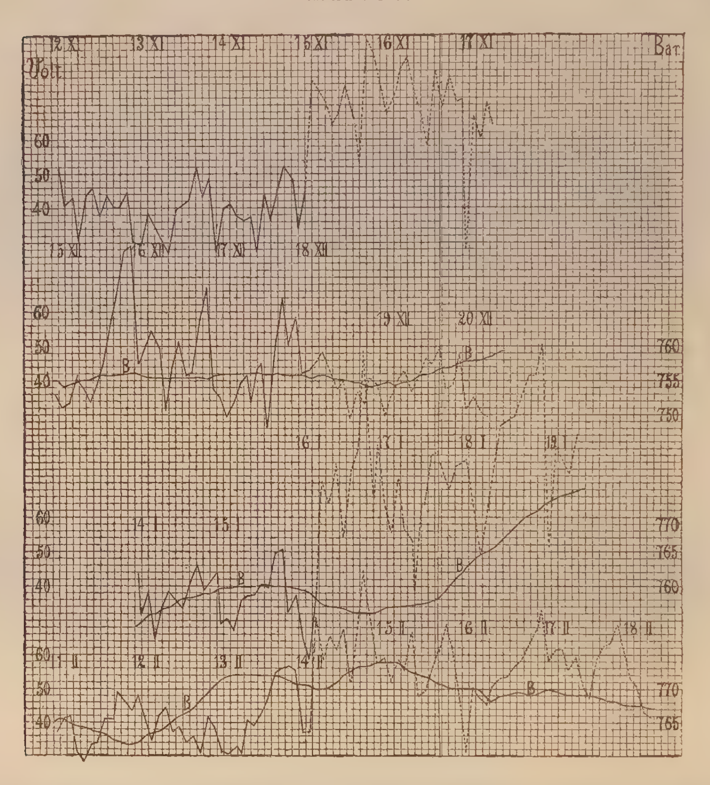
2) C. Dorno, Studie über Licht und Luft im Hochgebirge. Braunschweig 1911.

Temperaturschwankungen in den Versuchstagen im Versuchsraum.

Monat	Lichtversuche				Dunkelversuche			
	Temperatur Maximum	Temperatur Minimum	Differenz	Größte Schwankung innerhalb 24 Stunden	Temperatur Maximum	Temperatur Minimum	Differenz	Größte Schwankung innerhalb 24 Stunden
November	23,0	21,0	2,0	2,0	17,0	17,0	0	0
Dezember	17,0	16,35	0,65	0,45	15,90	15,05	0,85	0,5
Januar	18,0	15,9	2,1	0,7	13,90	13,40	0,5	0,5
Februar	19,6	18,9	0,7	0,45	17,0	16,80	0,2	0,2
März	20,80	19,45	1,35	1,0	18,0	17,60	0,4	0,4
April	22,85	21,0	1,85	1,40	20,30	19,80	0,5	0,4
Mai	21,2	19,7	1,5	0,95	19,20	18,50	0,7	0,35
Juni	17,05	16,50	0,55	0,20	15,40	14,40	1,0	0,5

von v. Schweidler[1]) auf Station Seeham u. a. m. Diese Messungen wurden aber meist mit Hilfe selbstregistrierender Instrumente und unter freiem Himmel angestellt, wobei das ganze System dem periodischen Wechsel verschiedener Außenfaktoren ausgesetzt war. Das hatte wohl zur Folge, daß die verschiedenartigsten Vermutungen aufgestellt wurden über die Bedeutung der einzelnen Außenfaktoren für das Entstehen der Rhythmik der Zerstreuung.

Kurve I.



Elektrizitätsverlust bei freistehendem Zerstreuungskörper innerhalb 5 Minuten im November, Dezember, Januar und Februar. Barometerkurve mit B bezeichnet. Das Datum, sowie die verstärkte Ordinatenlinie bezeichnet 12 Uhr mittags.

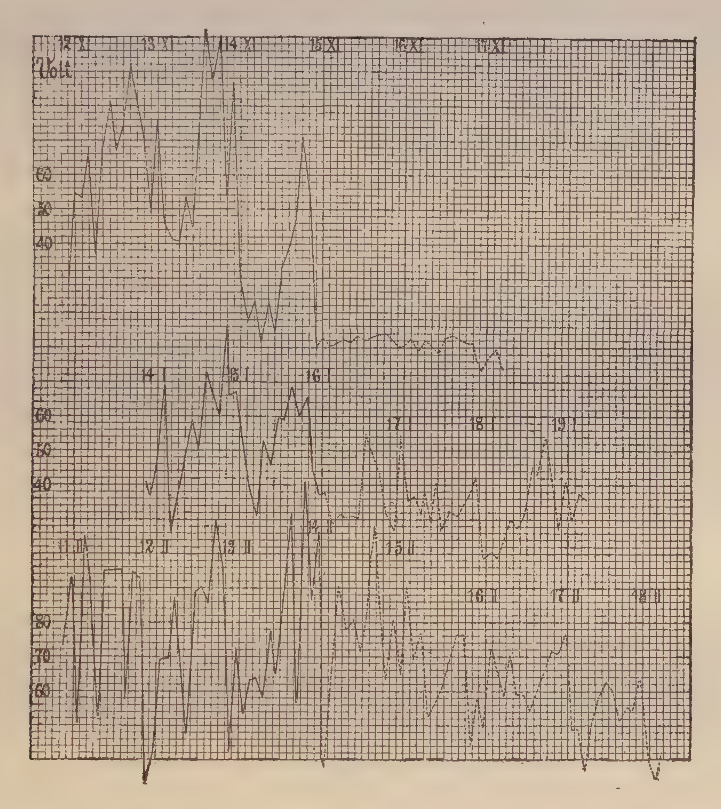
Es ist daher nicht unwesentlich, daß diese Rhythmik in hohem Maße auch in geschlossenen Räumen auftritt, wo in weit höherem

---

1) E. R. v. Schweidler, Beiträge zur Kenntnis der atmosphärischen Elektrizität XXV. Sitzungsber. d. Akademie d. Wiss. zu Wien 115. IIa. 1906. S. 1263—1271.

Grade eine Konstanz der Außenbedingungen herrscht. Es wäre daher dringend erwünscht, daß unter solchen Verhältnissen auch Versuche mit selbstregistrierenden Apparaten durchgeführt würden, damit die in meinen Angaben unvermeidlichen Fehler, die durch die direkte Beobachtung entstehen, sowie die Diskontinuierlichkeit der Angaben ausgeschaltet würden. Zudem übersteigt es die Kräfte eines Menschen, derartige Beobachtungen eine längere Zeit

Kurve II.



Elektrizitätsverlust im Jonenaspirator innerhalb 10 Minuten bei Versuchen im November, Januar und Februar. 12 Uhr mittags durch das Datum und die verstärkte Linie kenntlich.

hindurch Tag und Nacht durchzuführen. — Auf die in der Literatur verzeichneten Beobachtungen in geschlossenen Räumen will ich hier nicht näher eingehen, weil die Untersuchungen meist der Einheitlichkeit und konsequenten Durchführung ermangeln. Bald wurden die Messungen mit Anwendung eines Schutzzylinders aus-

geführt, bald ohne denselben. Mitunter wurde sogar die unzulässige Methode angewendet, die auf die eine Art gewonnenen Werte nach Umrechnung den auf die andre Art bestimmten gleichzusetzen. Ferner fehlen meist die Nachtbeobachtungen, und gerade diese sind für den Verlauf der Tageskurve am allerwichtigsten. — Wenn auch den in dieser Arbeit mitgeteilten Versuchen noch viele technische Fehler anhaften, so hoffe ich doch, daß sie als Grundlage für weitere wissenschaftliche Forschung nützlich sein werden.

Sehr undeutlich tritt die Rhythmik der Leitfähigkeitsintensitäten zu Tage bei der Dunkelkurve im März S. 410. Es war während der Versuchstage für Hamburger Verhältnisse ein besonders klarer Himmel, meist Sonnenschein, teils kalt. Ich muß es dahingestellt sein lassen, inwieweit die zufällige Kombination der meteorologischen Faktoren den abweichenden Typus dieser Kurve bedingt hat. Sie fällt den Ergebnissen der übrigen Monate gegenüber auf durch die sehr hohen, absoluten Zerstreuungswerte, und die großen Schwankungen derselben. Sie gleicht vielmehr den aus den Resultaten der Lichtbeobachtungen gewonnenen Kurven, die besonders im November, Januar, Februar ein sofortiges sehr starkes Ansteigen der Zerstreuung nach Belichtung des Versuchsraumes erkennen lassen. Die Periodizität ist während der Lichtperiode weniger ausgesprochen, die Schwankungen erheblich stärker. Es hat demnach den Anschein, als ob durch hohe Lichtzufuhr die Rhythmik der Leitfähigkeit überdeckt oder ausgelöscht wäre. Überhaupt konnte ich den Einfluß meteorologischer Faktoren wie Regen und Schneefälle, Gewitter und Klarheit der Luft oft beobachten. Bei den Lichtversuchen erhielt man sogar den Eindruck, als ob das Instrument diesen Einflüssen gegenüber „sensibler" wäre, als während der Dunkelheit.

Das starke Anwachsen der Zerstreuung vor Gewittern wurde schon von Gockel[1]), Zölß[2]), Thaller[3]), Kähler[4]) beobachtet, ebenso der Einfluß von Regengüssen und Schneefällen. Besonders häufig findet man aber die Transparenz der Luft als wirksamen

---

1) A. Gockel, Über den Jonengehalt der Atmosphäre. Meteorol. Zeitschr. 25, 1908, S. 9—19.

2) P. B. Zölß, Beiträge zur Kenntnis der atmosphärischen Elektrizität XVIII. Sitzungsber. d. Akademie d. Wiss. zu Wien. 114, IIa, 1905, S. 189—331.

3) R. Thaller, Beiträge zur Kenntnis der atmosphärischen Elektrizität L. Sitzungsber. d. Akademie d. Wiss. zu Wien 122, IIa, 1913, S. 1817—1828.

4) K. Kähler, Die Schwankungen des luftelektrischen Leitvermögens und des vertikalen Leitungsstromes am Erdboden. Ergeb. d. meteorol. Beobachtungen in Potsdam 1911, XVI—XXXI.

Faktor erwähnt. Diese Beobachtung bestätigte sich auch bei meinen Versuchen und zwar trat der Effekt so scharf und schnell hervor, daß es bedeutungslos zu sein schien, daß die Untersuchungen in einem Raum ausgeführt wurden, der von der Außenwelt rel. gut abgeschlossen war.

Die starke Steigerung der Zerstreuung während der Lichtperiode steht nicht im Einklang damit, daß bei den Dunkelkurven das Maximum gerade in der Nacht erreicht wurde. Es muß also nachts ein Faktor die Zerstreuung so stark beeinflussen, daß die Wirkung des Lichts am Tage nicht hervortritt. Freilich blieb auch diese Lichtwirkung bei einzelnen Versuchen, so im Dezember aus. Ich hatte derartig widersprechende Beobachtungen schon früher gemacht, als ich bei Tageslicht in einem Raum arbeitete, den ich lichtdicht verschließen konnte. Meist war auch hier der starke Rückgang der Zerstreuung nach Verdunkelung wahrzunehmen, aber nicht immer. Vielfach hatte ich schon große Differenzen in der Zerstreuungsgeschwindigkeit bei Versuchen, die sich gleich folgten und einmal in der Nähe des Fensters, dann in dem dunkleren Teil eines größeren Raumes ausgeführt wurden. Es ist auffallend, daß die Unterschiede bei den Licht- und Dunkelversuchen in den Sommermonaten überhaupt mehr verschwinden. — v. Schweidler[1]), der während einer partiellen Sonnenfinsternis Untersuchungen machte, konnte keine Veränderung der Zerstreuungswerte während der Verdunkelung bemerken. Diese Beobachtungen sprechen dafür, daß ein anderer Faktor die Lichtwirkung kompensieren oder anullieren kann.

Was die Beziehungen der Zerstreuung zum absoluten Stand des Barometers wie auch zu seinen Schwankungen anbetrifft, so ist bei meinen Versuchen von einem Einfluß des Luftdruckes nichts zu merken. Die von anderer Seite gemachten Angaben weichen in ihren Resultaten sehr von einander ab, Thaller[2]), der auch ohne Schutzdach arbeitete, fand eine Beziehung zwischen dem Barometerstand und der Zerstreuung, dagegen fehlte ein merkbarer Einfluß der Barometerschwankungen auf den Elektrizitätsverlust. Nach Schweidler[3]) ist (S. 1065) „der Gang des Luftdruckes von viel geringerem Einfluß, als der Absolutwert, und

1) E. R. v. Schweidler, Beiträge zur Kenntnis der atmosphärischen Elektrizität XXII. Sitzungsber. d. Akad. d. Wiss. zu Wien 114, IIa, 1905, S. 1705—1735.

2) R. Thaller, Beiträge zur Kenntnis der atmosphärischen Elektrizität L. Sitzungsber. d. Akademie d. Wiss. zu Wien, 122, IIa, 1913, S. 1817—1828.

3) E. R. v. Schweidler, Beiträge zur Kenntnis der atmosphärischen Elektrizität LIII. Sitzungsber. d. Akad. d. Wiss. zu Wien, 125, IIa, 1916, S. 997—1082.

zwar ist die Leitfähigkeit bei steigendem Luftdruck unbedeutend erhöht, bei fallendem erniedrigt — also umgekehrt, als es zu erwarten wäre, falls man in der Bodenatmung d. i. dem Austreten emanationshaltiger und ionenreicher Bodenluft in die freie Atmosphäre die Ursache setzen wollte". Auch Mache[1]) sagt S. 1380 „Als solches erscheint mir doch der Schluß, daß die Luftdruckschwankungen allein schwerlich im Stande sind, den Emanationsverbrauch der Erde zu decken". Nach Conrad[2]) hat das Steigen und Sinken des Luftdruckes auf dem Sonnblick keinen nachweisbaren Einfluß auf die Zerstreuung, doch fand er bei höherem Luftdruck größere Zerstreuungswerte. Mazelle[3]) konnte keine Abhängigkeit der Zerstreuung weder vom absoluten Luftdruck noch von seinen Schwankungen erkennen. Kohlrausch[4]) wiederum fand bei fallendem Barometer größere Jonenzahlen und eine gesteigerte Leitfähigkeit und Müns[2]) sagt S. 34: „Eine gesetzmäßige Beeinflussung der einzelnen luftelektrischen Werte, außer Q, durch den Barometerstand ist unverkennbar. Überall finden wir kleine Werte bei niedrigem Barometerstande. Bei normalem Luftdruck 760 ± 5 mm werden alle Werte größer, um bei ungewöhnlich hohem Luftdruck wieder kleiner zu werden". Und S. 35 sagt er weiter: „Die Luftdruckschwankungen machen sich bei den einzelnen Werten nicht so hemerkbar, wie es zu erwarten wäre". Zölß[6]) konnte bei seinen Untersuchungen in den Kellerräumen der Sternwarte zu Kremsmünster keine Beziehungen weder des Standes des Barometers noch seiner Schwankungen zu der Leitfähigkeit der Atmosphäre finden, während er in der Steinbruchhöhle eine sehr deutliche mit den Schwankungen des Barometers feststellte, derart, daß bei sinkendem Luftdruck die Leitfähigkeit stieg und umgekehrt. Das negative Ergebnis bei den Untersuchungen im Keller schiebt Zölß auf den Steinfußboden des

---

1) H. Mache, Beiträge zur Kenntnis der atmosphärischen Elektrizität, XXI. Sitzungsber. d. Akademie d. Wiss. zu Wien, 114, IIa, 1905, S. 1377—1388.

2) V. Conrad, Beiträge zur Kenntnis der atmosphärischen Elektrizität, XIX. Sitzungsber. d. Akademie d. Wiss. zu Wien. 114, IIa, 1905, S. 335—353.

3) E. Mazelle, Beiträge zur Kenntnis der atmosphärischen Elektrizität, XX. Sitzungsber. der Akademie d. Wiss. zu Wien, 114, IIa, 1905, S. 399—501.

4) K. M. F. Kohlrausch, Beiträge zur Kenntnis der atmosphärischen Elektrizität, LII. Sitzungsber. d. Akademie d. Wiss. zu Wien, 123, IIa, 1914, S. 2321—2348.

5) Rudolf Müns, Über Jonisation und Radioaktivität der Atmosphäre in Kiel. Dissertation. Kiel 1913.

6) P. B. Zölß, Beiträge zur Kenntnis der atmosphärischen Elektrizität. XVIII. Sitzungsber. d. Akademie d. Wiss. zu Wien, 114, IIa, 1905, S. 189—331.

Raums. Dennoch kommt er S. 231 zu dem Schluß: „Die Wirkung der Luftdruckschwankungen reicht nicht hin, die tägliche Periode der Elektrizitätszerstreuung zu erklären“. Auch in den zusammenfassenden Arbeiten über Luftelektrizität von Gockel[1]), Mache und Schweidler[2]) und Kähler[3]) wird der Einfluß des Luftdruckes als wenig entscheidend für den Gang der Zerstreuung hingestellt.

Diese Literaturangaben, die noch durch weitere sich widersprechende zu vermehren wären, genügen wohl schon, um zu beweisen, daß die Bodenatmung keine hinreichende Erklärung ist für die Schwankungen der Jonisation, oder aber, daß die Bodenatmung in erster Linie von anderen Faktoren als vom Luftdruck abhängig ist. Auch Müns[4]) meint S. 45, daß seine Ergebnisse keine Stütze sind für die Ebertsche Theorie und Zölß[5]) deutet S. 215 an, daß noch eine andere Ursache als die Luftdruckschwankungen die Änderungen des Emanationsgehaltes der Atmosphäre bedingen müssen. Gockel[6]) sagt S. 63: „Die geringe Übereinstimmung in den anderen Fällen beweist entweder, daß auch noch andere Quellen der Jonisation existieren als die radioaktive Emanation, oder daß die Schwankungen der Produktion von Jonen durch die Emanation kompensiert werden können durch andere Umstände, welche einerseits eine Ansammlung von Jonen, andererseits eine Vernichtung derselben durch Adsorption, Molisierung usw. zur Folge haben. Meines Erachtens treten sowohl die letzteren Umstände auf, als auch: Die Emanation ist nicht die einzige Quelle der Jonisation“. Am naheliegendsten ist es, die durchdringende Strahlung für die Schwankungen der Jonisation als ausschlaggebenden Faktor heran zu ziehen. Diese Erscheinung kann aber nicht in Betracht kommen, denn wenn überhaupt eine Tagesperiode der durchdringenden Strahlung besteht, so sind die Differenzen viel zu gering, um den beobachteten Effekt bei der Leitfähigkeit bedingen zu können, zudem sprechen die bisherigen

---

1) A. Gockel, Die Luftelektrizität. Leipzig 1908.

2) H. Mache und E. v. Schweidler, Die atmosphärische Elektrizität. Die Wissenschaft. Heft 30. 1909.

3) K. Kähler, Luftelektrizität. Sammlung Göschen 1913.

4) Rudolf Müns, Über Jonisation und Radioaktivität der Atmosphäre in Kiel. Dissertation. Kiel 1913.

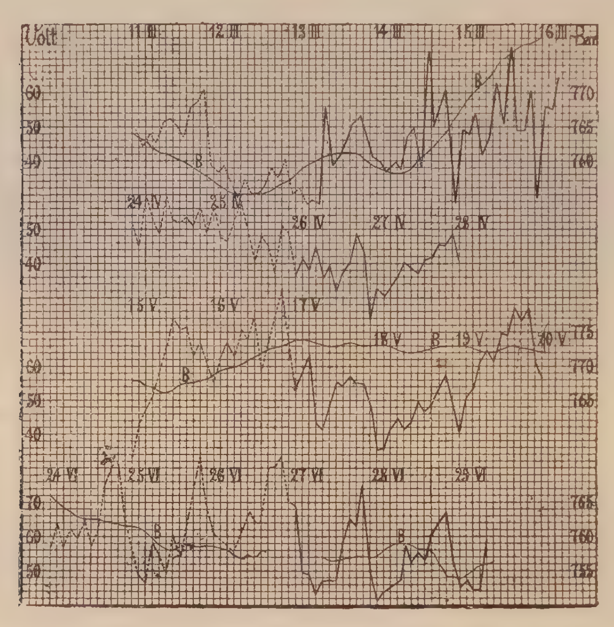
5) P. B. Zölß, Beiträge zur Kenntnis der atmosphärischen Elektrizität, XXIII. Sitzungsber. d. Akademie d. Wiss. zu Wien, 114, IIa, 1905, S. 189—331.

6) A. Gockel, Über den Jonengehalt der Atmosphäre. Meteorolog. Zeitschr. 23, 1906, S. 53—67.

Angaben mehr dafür, daß gerade das Minimum der durchdringenden Strahlung in die Tageszeit fällt, wo die Zerstreuung ihr Maximum aufweist. Sie würde also eher in dem entgegengesetzten Sinne einwirken.

Da unter den bei meinen Versuchen herrschenden Umständen weder die Schwankungen der Temperatur noch der Feuchtigkeit den rhythmischen Gang der Zerstreuung verursacht haben können,

Kurve III.



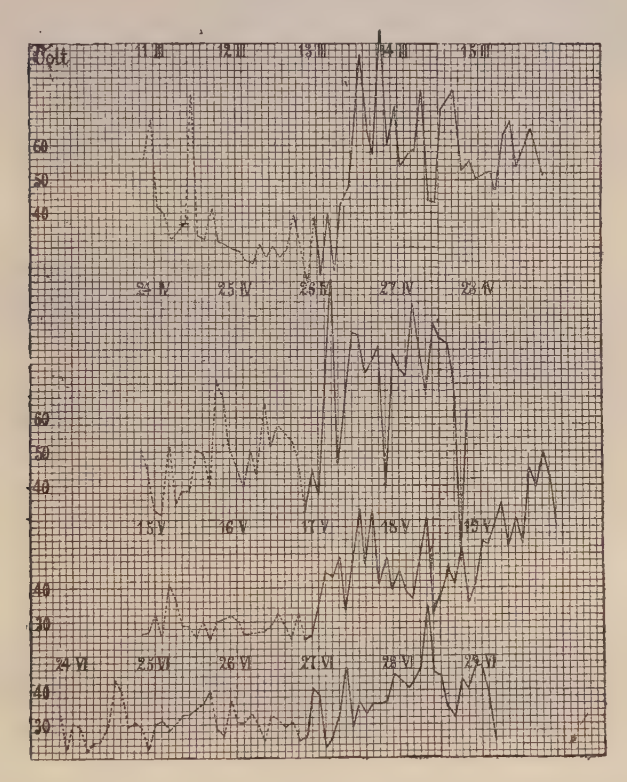
Elektrizitätsverlust bei freistehendem Zerstreuungskörper innerhalb von 5 Minuten bei + geladener Elektrode bei Versuchen im März, April, Mai und Juni. Barometerkurve mit B bezeichnet.

der Luftdruck ohne wesentlichen Einfluß zu sein scheint und das Licht des Tageswechsels sogar im entgegengesetzten Sinne wirken müßte, als es der Tageskurve entspricht, so müssen wir zugeben, daß die Ursache der Rhythmik der Zerstreuung noch ganz ungeklärt ist, und daß wir hier möglicherweise noch einen unbekannten kosmischen Faktor oder eine unbekannte Kombination von Faktoren vor uns haben, die vielleicht auch für rhythmische biologische Erscheinungen (Schlaf) von dem maßgebendsten Einfluß ist. Jedenfalls liegt hier noch ein weites Feld für den Forscher offen, und die ausstehenden Resultate können auch biologisch von der weitgehendsten Bedeutung sein.

Ich gehe nun über zu der Besprechung der Resultate, die ich mit dem Jonenaspirator erhalten habe s. S. 405 u. 411. Auch hier tritt bei den Dunkelversuchen bisweilen eine Tagesperiodizität auf, so im November, Januar und Februar S. 405. Die Dezemberversuche mußten leider wegen eines Federbruches im Uhrwerk ausfallen.

Die zu beobachtende Periodizität erscheint nicht so ausgesprochen, wie es bei dem freistehenden Zerstreuungskörper der


Kurve IV.



Elektrizitätsverlust im Jonenaspirator bei + geladener Elektrode innerhalb 10 Minuten bei Versuchen im März, April, Mai und Juni.

Fall war, auch kann von einem Parallelismus der Leitfähigkeits- und Jonisationskurve höchstens im November und Januar und nur in sehr beschränktem Maß die Rede sein.

Das Licht wirkt auf den Jonengehalt gerade im dem entgegengesetzten Sinn, wie auf die Leitfähigkeit. Während der Dunkelperiode sind nicht allein die absoluten Werte der Jonisation

erheblich höher, sondern auch die Schwankungen sind wesentlich größer als im Licht. Diese einengende Wirkung des Lichtes nur auf die Veränderungen von Temperatur und Feuchtigkeit zu setzen, scheint unangängig, weil die Temperatur während der Dunkelversuche im April fast übereinstimmte mit den Temperaturen der Lichtperiode im März. Die Jonisation ist aber dennoch im April eine sehr viel größere als im März. Auch aus einem kleinen Nebenversuch geht hervor, daß das Licht noch in anderer Hinsicht von Bedeutung sein muß. Ich stellte nämlich im November vom 16ten abends 8 Uhr ab die Lampen so hintereinander, daß sie in einer Linie mit der Achse des Jonenaspirators standen aber in möglichst der gleichen Entfernung von demselben wie vordem. Der Lichtschein fiel infolgedessen in das Rohr des Aspirators hinein und beleuchtete die innere Elektrode und den inneren Teil der äußeren Elektrode eine Strecke weit. Sofort sanken dann die ohnehin schon sehr niedrigen Novemberwerte noch um ein Beträchtliches, so daß nicht einmal mehr ein Abfall von 10 Volt in 10 Minuten erreicht wurde. Es ist auffallend, daß beim Jonenaspirator die direkte Belichtung der Innenelektrode die Elektrizitätsabgabe derselben herabsetzt, beim freistehenden Zerstreuungskörper Belichtung dagegen die Elektrizitätsabgabe steigert. Der Unterschied, daß im ersteren Fall ein Sättigungsstrom besteht, im letzten aber in erster Linie ein freier Strom, dürfte kaum für die Erklärung der Beobachtung von Bedeutung sein.

In wie weit die übrigen meteorologischen Faktoren wie die Schwankungen von Temperatur, Feuchtigkeit, Luftdruck sowie Niederschläge, Gewitter und Transparenz der Luft im Stande sind, die großen Schwankungen der Jonisation zu erklären, muß ich dahingestellt sein lassen. Auch hier ist noch manche offene Frage, doch genügen meine Versuche nicht, um weiter auf dieselben einzugehen. —

Zum Schluß sei mir noch gestattet, die Kurven der Leitfähigkeit und des Jonengehaltes in den entsprechenden Monaten mit einander zu vergleichen. Es fällt dabei nicht nur die gerade gegensätzliche Wirkung des Lichtes in bezug auf die absolute Höhe der Abfallswerte als auch dem Grad der Schwankungen gegenüber den Dunkelkurven auf, sondern das Steigen und Sinken der Kurven des Jonenaspirators geht bisweilen parallel, bisweilen der Leitfähigkeitskurve entgegengesetzt, meist jedoch scheint keine Beziehung der beiden Kurven zu einander zu existieren.

Da die Leitfähigkeit der Atmosphäre eine Funktion ist der Jonenzahl und der Jonenbeweglichkeit, so kann uns eine Zusammen-

stellung der Kurven über Leitfähigkeit und Jonisation ein Bild geben von den Veränderungen der Jonenbeweglichkeit. Nach Schering[1][2]) ist es erlaubt, die Kurve, die sich aus den Messungen mit freiem Zerstreuungskörper ergeben, als nicht erheblich abweichend von der Kurve der Leitfähigkeit anzusehen. Die Kurve des Jonenaspirators ist ein Bild der Schwankungen der Jonenzahlen. Eine entsprechende Verwertung beider Kurven muß daher ein Maß für die jeweiligen Jonengeschwindigkeiten abgeben. — Es ist bekannt, daß diese Geschwindigkeiten Veränderungen unterworfen sind.

Vergleichen wir die Kurven der Leitfähigkeit und Jonenmenge z. B. vom November o. S. 404 u. 405. In der Dunkelperiode ist die Leitfähigkeit relativ klein, die Jonenmenge dagegen groß, also muß eine relativ kleine Jonengeschwindigkeit geherrscht haben. — In der Lichtperiode desselben Monats ist die Leitfähigkeit groß, die Jonenmenge sehr klein, diese Werte entsprechen demnach besonders großen Jonengeschwindigkeiten. Nun stellte aber schon Kohlrausch[3]) fest, daß die Jonengeschwindigkeit mit der wachsenden Feuchtigkeit zunimmt, und Müns[4]), der in der Aktivierungszahl A ein ungefähres Maß der Jonenbeweglichkeit sieht, sagt dennoch S. 37: „Ganz auffallend sind die kleinen Werte der Jonenbeweglichkeiten bei trockener und die großen Jonenbeweglichkeiten bei feuchter Luft. Gerade das Umgekehrte hätte man erwarten sollen". Wie stimmen nun diese Ergebnisse mit meinen Beobachtungen? Stimmen meine Voraussetzungen mit den tatsächlichen Verhältnissen auch nur einigermaßen überein, dann hätte ich kleine Jonenbeweglichkeiten während der Dunkelperiode, sehr große im Licht. Es unterliegt aber keinem Zweifel, daß in der Lichtperiode die Feuchtigkeit der Atmosphäre eine geringere war als in der Dunkelheit. Nach Müns[5]) und Kohlrausch[6]) wären

1) H. Schering, Der Elster-Geitelsche Zerstreuungsapparat und ein Versuch quantitativer absoluter Zerstreuungsmessung. Dissertation. Göttingen 1904.

2) H. Schering, Der Elster-Geitelsche Zerstreuungsapparat und ein Versuch quantitativer absoluter Zerstreuungsmessung. Ann. d. Physik, 20. 1906, S. 174—195.

3) K. M. F. Kohlrausch, Beiträge zur Kenntnis der atmosphärischen Elektrizität LII. Sitzungsber. d. Akademie d. Wiss. zu Wien, 123, IIa, 1914, S. 2321—2348.

4) Rudolf Müns, Über Jonisation und Radioaktivität der Atmosphäre in Kiel. Dissertation. Kiel 1913.

5) Ebd.

6) K. M. F. Kohlrausch, Beiträge zur Kenntnis der atmosphärischen Elektrizität LII. Sitzungsber. d. Akademie d. Wiss. zu Wien, 123, IIa, 1914 S.2321—2348.

aber gerade die umgekehrten Resultate zu erwarten gewesen. Durch die Annahme der Wiedervereinigung von Jonen ist der Widerspruch auch nicht so einfach zu beseitigen. Also auch in dieser Beziehung sind Unstimmigkeiten, die einer gründlichen Bearbeitung bedürften.

Erstaunlicherweise sind Parallelversuche mit Jonenaspirator und freistehendem Zerstreuungskörper wie die mitgeteilten bisher scheinbar außerordentlich selten und nur mehr gelegentlich angestellt worden. Aber nicht nur das, sondern die Untersuchungen der Luftelektrizität unter Bedingungen, deren Konstanz bis zu einem gewissen Grad in der Gewalt des Experimentators stehen, sind bislang überhaupt noch sehr spärlich. Diese Unterlassungssünden mögen in erster Linie Schuld daran sein, daß für die Schwankungen der Leitfähigkeit Faktoren verantwortlich gemacht wurden, die jedenfalls nicht in erster Linie bedingend sind. — Möge nun diese Arbeit trotz ihrer mancherlei Unzulänglichkeiten dazu beitragen, daß den angeschnittenen Fragen ein weiteres Interesse entgegengebracht wird. Geschieht das, so sehe ich die erhebliche aufgewendete Arbeit, die mit den positiven Ergebnissen der Untersuchungen nicht im Einklang steht, dennoch als gut angewendet an. Ich hoffe, daß ich mich selbst dem Studium der Leitfähigkeit in geschlossenen Räumen bei einwandfreier Versuchsanstellung noch ferner werde zuwenden können.

Zusammenfassung der Resultate.

Es wurden Bestimmungen der Leitfähigkeit und des Jonengehaltes der Atmosphäre gemacht im geschlossenen Raum vom November 1917 bis Juni 1918, in jedem Monat 5—8 Tage lang bei etwa 2stündlichen Ablesungen in dauerndem Licht und dauernder Dunkelheit, bei + Ladung der Elektroden.

Die Intensität der Leitfähigkeit unterlag bei dauernder Dunkelheit einem tagesperiodischen Wechsel. In den ersten Morgenstunden — zwischen 2 und 4 Uhr meistens — ist die Zerstreuung am größten, im Lauf des Tages wurde der niedrigste Wert erreicht.

Während der Lichtperiode war die Periodizität der Leitfähigkeit meist weniger ausgesprochen, die Schwankungen größer, und die Intensität besonders in den Wintermonaten eine erheblich gesteigerte.

Weder die Schwankungen des Luftdruckes noch der Temperatur und Feuchtigkeit geben eine Erklärung für den rhythmischen Gang der Leitfähigkeitskurve.

Die Jonenmenge zeigte in den Dunkelperioden im Winter ebenfalls einen tagesperiodischen rhythmischen Wechsel der Intensität. Im Sommer wurden die Kurven unregelmäßiger.

Belichtung wirkt auf die Jonenmenge gerade umgekehrt wie auf die Leitfähigkeit. Der Elektrizitätsverlust im Jonenaspirator war während der Lichtperiode erheblich niedriger, die Schwankungen geringer.

Die Veränderung der Jonenbeweglichkeit durch das Licht muß sehr erheblich sein, um die gegensätzlichen Resultate bei der Bestimmung der Leitfähigkeit und gleichzeitig des Jonengehaltes mit einander in Einklang zu bringen.

Hamburg, Jan. 1919.

Insitut für Allgemeine Botanik.

Feuerschlagen.

Von

A. Johnsen.

(Mineralog. Institut der Universität Kiel.)

Vorgelegt von O. Mügge in der Sitzung vom 16. Mai 1919.

I. Stahlfunken.

Daß diejenigen Mineralien, welche härter als Stahl sind, mit ihm zusammengeschlagen, Funken liefern, hoben schon C. v. Linne[1]) und A. G. Werner[2]) hervor; die Härte der gewöhnlichen Stahlsorten, die man zu Messerklingen, Feilen und Feuerzeugen verwendet, liegt etwa bei H = 5.5 der Mohs'schen Skala. Daher funken am Stahl beispielsweise Quarz (H = 7), Flint[3]) (H = 7), Schwefelkies[3]) (H = 6.5), Markasit (H = 6), Arsenkies (H = 5.5 bis 6), Magnetit (H = 6), Eisenglanz (H = 6). In allen diesen Fällen werden Stahlteilchen durch den Stoß erhitzt und zugleich, bevor die Wärme auf die übrige Stahlmasse abfließen und dadurch die Temperatur sinken

---

1) C. v. Linné, „Natursyst. d. Mineralreichs", deutsch von J. Fr. Gmelin. I. S. 59. Nürnberg 1777.

2) A. G. Werner, „Von den äußerlichen Kennzeichen der Fossilien" S. 249, Leipzig 1774. Bei G. Agricola ist dieser Gedanke noch nicht ganz ausgeprägt, denn er sagt („De ortu et causis subterraneorum", liber IV. Cap. XVI. p. 120 Wittebergae 1613), besonders hart seien diejenigen Mineralien, die an Stahl oder aneinander geschlagen Funken gäben.

3) Im sog. „Nürnberger Radschloß" der Büchsen des 16. Jahrhunderts drehte sich ein Stahlrad gegen einen Flint oder gegen einen Schwefelkies. Dieser heißt daher noch in J. Fr. Henckel's „Pyritologia" oder „Kieshistorie" (S. 72. Leipzig 1725) „Büchsenstein", während vom niederdeutschen „Flint" das Wort „Flinte" stammt.

kann, losgerissen; jene Erhitzung leitet in den Stahlpartikeln eine Verbrennung ein, welche Gelbweißglut erzeugt und Eisenoxyde entstehen läßt.

II. Mineralfunken.

Wie in modernen Feuerzeugen der Stahl von dem weicheren Cereisen Teilchen losreißt, die dann verglimmen, so können gewisse Mineralien, an einem härteren Mineral oder, wenn sie weicher als Stahl sind, an diesem oder endlich auch an ihresgleichen geschlagen, Funken hergeben. Hierbei muß das funkende Mineral erstens seiner chemischen Natur nach einer Oxydation fähig und zweitens genügend hart sein, sodaß die offenbar mit abnehmender Härte abnehmende Reibungswärme zur Einleitung der Oxydation ausreicht. So funken an Quarz und Flint (H = 7) sowie an ihresgleichen der Schwefelkies (H = 6,5) und der Markasit (H = 6), beide von der Formel FeS_2, unter Verbreitung des Geruches und Schwefeldioxyd. Wahrscheinlich verbrennt hierbei, wie nach Berzelius beim Rösten des Schwefelkieses, nur eines der beiden Schwefelatome. Freilich funkt an Quarz und Flint auch der Magnetkies (FeS; H = 4) von Bodenmais und, wenngleich weniger, künstliches Schwefeleisen, doch ergeben die chemischen Analysen von Magnetkies einen Schwefelüberschuß bis über 3 % und somit vielleicht einen Schwefelkiesgehalt bis über 18,5 %. Eine Feststellung, ob auch der Troilit (FeS) der Meteoriten, der keinen Schwefelüberschuß besitzt, Funken hergibt, wäre entscheidend.

Die rötlichgelb glühenden Schwefelkiesfunken fallen z. T. bis 1 m tief, bevor sie erlöschen; ihre gleichförmige Fallgeschwindigkeit beträgt etwa 50 cm/sec., was einen Funkendurchmesser von ungefähr $^1/_{20}$ mm ergibt, falls man die Stokes'sche Widerstandsformel hier anwenden darf.

Dagegen scheinen keine Funken zu liefern Kobaltglanz (CoAsS; H = 5.5), Gaukodot ((Fe, Co) AsS; H = 5), Speiskobalt ($CoAs_2$; H = 4.5 bis 5), Kupferkies ($CuFeS_2$; H = 3.5 bis 4), Buntkupfererz (Cu_3FeS_3; H = 3), Kupferglanz (Cu_2S; H = 2.5 bis 3) und Zinkblende (ZnS; H = 3.5 bis 4); auch Arsenkies (FeAsS; H = 5.5 bis 6) funkt mit einem härteren Mineral zusammengeschlagen, nicht, obwohl sich dabei Arsen unter Entwicklung des bekannten knoblauchartigen Geruches oxydiert.

III. Flintpartikeln.

Stößt oder reibt man zwei Flintstücke aneinander, so sieht man im Dunkeln und sogar bei Tageslicht an den Berührungsstellen hellgelbe Lichtpünktchen aufblitzen. Diese am Flint wie am Achat und am Quarz wohl zuerst von J. H. Pott[1]) erkannte kalte Strahlung zeigt sich ebenso unter Wasser wie an Luft und heißt Tribolumineszenz oder besser „Frakturolumineszenz“, denn es wird nicht etwa durch die Reibungswärme eine sog. Thermolumineszenz[2]) erregt, sondern durch das Abbrechen von kleineren oder größeren Bruchstücken entsteht ein Ausgleich vorheriger elektrischer Ladungen[3]), der die beteiligten Elektronen in Schwingung versetzt. Diese erlischt binnen $^1/_{100}$ bis $^1/_{10}$ Sekunde nach dem Abspringen der Partikeln, sodaß man letztere, kaum bemerkt; bei solch kurzer Erregungsdauer wie beim Zerbrechen treten nach P. Lenard[4]) nur die „Momentanbanden“ der Lumineszenz statt der „Dauerbanden“ der Fosforeszenz auf.

Der von A. Imhof[5]) für lumineszierende Quarzteilchen festgestellte „Minimalkorn“-Durchmesser von 0,1 mm im Mittel dürfte auch für den verwandten Flint gelten, während er bei Flußspat nahezu 1 mm beträgt. Dieses Mineral zeigt, ebenso wie Zinkblende, nicht in allen Varietäten gleich starke Frakturolumineszenz, dagegen scheint sie allen Flintvorkommen und Achaten, allen Quarzen (pyrogenen wie hydatogenen), allen Orthoklassen (Sanidinen wie Adularen), allen Obsidianen und den böhmischen Moldawiten (im Gegensatz zu künstlichen Gläsern) in (nahezu gleich) hohem Grade eigentümlich zu sein.

Auch Körper, die weniger hart als Flint, Quarz etc. sind,

---

1) J. H. Pott, „Chymische Untersuchungen etc.“ S. 39. Potsdam 1746. Ähnliches Verhalten beschrieb A. Imhof (Fysikal. Zeitschr. **18**. S. 78 1917) an Opal und Quarzglas.

2) Das von Wall (Philos. Transact. **26**. p. 73, London 1710) entdeckte Leuchten von Diamant beim Reiben mit Holz, Wolle u. dergl. ist, da hierbei kaum Partikeln abspringen werden, Thermolumineszenz; dagegen ist das beim Abspringen schadhafter Stellen während des sog. Graumachens von Diamant auftretende Leuchten Frakturolumineszenz.

3) Daher zeigen nur Dielektrika jene Lumineszenzen. Die vorherige Ladung kann lichtelektrische (Lenard) oder piezoelektrische (Voigt u. a.) Ursachen haben.

4) P. Lenard, Sitzgsber. Heidelb. Akad. Wiss. Abh. 8. S. 33. Heidelberg 1909.

5) A. Imhof, Fysikal. Zeitschr. **18**. S. 78. 1917.

können deren Lumineszenz erregen, wofern sie kleine Zacken von diesen Mineralien abzusprengen vermögen.

Unabhängig von jenem Leuchten müssen sich die abspringenden Flintpartikeln durch Stoß, Reibung und Abbrechen erwärmt haben. Diese Partikeln entzünden reine trockne Watte nicht, wohl aber solche, die mit Schwefelkohlenstoff getränkt ist. Da man diese Entflammung auch mit schmelzendem Zink, nicht aber mit schmelzendem Blei erzielen kann und deren Schmelzpunkte bei 419° und 327° liegen, so sind einzelne Flintteilchen jedenfalls auf über 327° erhitzt; andererseits kann ihre Temperatur nicht oberhalb 500° liegen, da man sonst rotglühende Funken wahrnehmen würde.

Durch Zusammenschlagen zweier gleicher Mineralien wird eine Entzündung umso schneller bewirkt werden können, je heißer und je zahlreicher die abfliegenden Partikeln sind. Die Erhitzung der Teilchen wächst, von *verbrennenden* Mineralien wie Schwefelkies etc. abgesehen, mit der Härte, ihre Anzahl aber nimmt mit zunehmender Härte ab; daher muß ein Optimum der Zündfähigkeit existieren, welches bei der Härte $H = 7$ von Flint und Quarz zu liegen scheint. Die spezifischen Wärmen der verschiedenen Mineralien haben wenig Einfluß auf deren Zündfähigkeit, da sie nur um etwa 15 % zu differieren pflegen. Auch das Wärmeleitungsvermögen spielt kaumeine Rolle, weil die Erhitzung der abspringenden Partikeln nahezu adiabatisch erfolgt.

Die Reibungswärme beim Aneinanderschlagen oder Wetzen eines Flintstückes an seinesgleichen oder an einem andern genügend harten Material hat oft noch eine weitere Folge: es entsteht häufig der Geruch nach verbrennendem Horn. Dieser tritt aber an solchen Flintscherben, die soeben frisch aus einer Knolle herausgeschlagen sind, *nur dann* auf, wenn man sie zunächst einige Sekunden lang, etwa wie ein Stück Seife beim Waschen, zwischen den Händen gerieben hat. Die Epidermisschüppchen der Hand, die hierbei am Flint haften geblieben sind, versengen beim Aneinanderschlagen infolge der Erhitzung, da Horn schon unterhalb 200° verkohlt. Dieser Horngeruch muß daher bei allen denjenigen Mineralien auftreten, die rauh genug zum Erfassen jener Hornhautteilchen und hart genug zur Entwicklung der nötigen Reibungswärme sind. Diese Folgerung wird u. a. von den rauhen Korallendolomiten, an denen schon G. *Agricola*[1])

---

1) G. *Agricola* l. c. liber IV. Cap. XV. p. 118.

und C. Gesner[1]) jenen Geruch beim Wetzen wahrnahmen, sowie durch die matten Quarzgeschiebe der Flüsse vollauf bestätigt, während Bergkristall zu glatt und Alabaster zu weich ist.

IV. Feuergewinnung der Steinzeitmenschen.

L. Hopf[2]) berichtet, daß es ihm, allerdings nach langer Bemühung, gelungen sei, mit zwei Flintstücken Feuer zu schlagen, und erwähnt ähnliche Angaben von Vergil, Seneca und Plinius, spricht sich indeß über die Natur des feuerfangenden Stoffes nicht aus. Die obigen Temperaturschätzungen dürften aber beweisen, daß jedenfalls die den Steinzeitmenschen bekannten Stoffe wie Zunderschwamm, Moos und dergl. schwerlich mittels zweier Flinte zur Entflammung gebracht werden können.

So sagt auch H. Obermaier[3]), daß die Kunst Feuer zu schlagen etwas jünger sei als diejenige, mit Holz Feuer zu reiben oder zu sägen, weil man in ersterem Falle nur dann Erfolg zu haben pflege, wenn außer Flint und Zunder auch Schwefelkies benutzt werde.

Ferner erzählt F. Ratzel[4]), daß die Pescheräh sich Feuer durch Gegeneinanderschlagen „eisenhaltiger Kiesel" bereiteten, die sie von der Insel Clarence importierten. Der besondere Import von diesem Eiland nach dem Feuerland spricht wiederum dafür, daß die überall verbreiteten gewöhnlichen Kiesel (Quarz und Flint) nicht geeignet sind. Unter Ratzel's „eisenhaltigen Kieseln" hat man vermutlich Quarz mit eingesprengtem Schwefelkies oder gar diesen allein zu verstehen.

Wenn vorgeschichtliche Menschen sich mit Flint Feuer bereitet hätten, würden wir unter ihren Grabbeigaben wahrscheinlich zwei Flintscherben finden; das ist aber bisher nie der Fall gewesen. Dagegen wurde sowohl in der nordischen Ganggräberzeit (jüngere Steinzeit) als auch noch in der nordischen Bronzeperiode den Toten zuweilen eine Flintscherbe nebst einer — inzwischen stark zersetzten — Schwefelkiesknolle[5]) mitgegeben.

---

1) C. Gesner, „De omni rerum fossilium genere etc." p. 37 u. 53 Tiguri 1565. Das dort als Hildesheimer Marmor bezeichnete Gestein kann nach einer freundl. Auskunft des Herrn Privatdoz. Dr. Wetzel hier nur Korallendolomit sein, und dieser verbreitet in der Tat einen intensiven Geruch der beschriebenen Art.

2) L. Hopf, Zeitschr. Kosmos S. 304, 1917.

3) H. Obermaier, „Der Mensch der Vorzeit" S. 418. 1912.

4) F. Ratzel, „Völkerkunde" I. S. 533. 2. Aufl. 1894.

5) Derartige Grabbeigaben wurden mir von Direktor Dr. Knorr und Kustos Rothmann im „Schleswig-Holsteinischen Museum vaterländischer Altertümer" zu Kiel gütigst gezeigt.

So nennen denn auch Plinius und Dioskorides (1. Jahrh. n. Chr.) den Schwefelkies „Pyrites“ d. h. „Feuerstein“.

Daß die unter Gelbglut verbrennenden Schwefelkiesteilchen, die wir oben beschrieben, ungefähr ebenso leicht wie Stahlfunken die Entflammung von Zunder und dergl. herbeiführen können, ist selbstverständlich; in beiden Fällen ist die Verbrennungswärme der Gramm-Molekel von der Größenordnung 10^4 bis 10^5 gcal.

Das Zustandsdiagramm von Kalium- Natriumchlorid und ihre Ätzfiguren.

Von

G. Tammann.

Vorgelegt in der Sitzung vom 4. Juli 1919.

Kalium- und Natriumchlorid bilden nach ihrer Kristallisation aus ihren flüssigen Mischungen eine lückenlose Mischkristallreihe[1]. Die Schmelzkurve hat ein schwaches Minimum bei 664° und 0,5 Mol KCl. Zwischen 400 und 300° findet die Entmischung der Mischkristalle unter ziemlich erheblicher Wärmeentwicklung statt. Die Temperaturen des Beginns dieses Vorganges liegen auf einer Kurve, die bei 406° und 0,33 Mol KCl ihr Maximum hat. Bei schneller Abkühlung läßt sich diese Entmischung unterdrücken, die Mischkristalle also auch bei tieferen Temperaturen realisieren. R. Nacken[2] verfolgte optisch diesen Entmischungsvorgang zwischen 0.1 bis 0.9 Mol KCl und fand, daß die Entmischungskurve von etwa 300° bis zu einem Maximum bei 495° und etwa 0.35 Mol KCl ansteigt. Dementsprechend ist der Wärmeinhalt der abgeschreckten Schmelzen größer als der der langsam gekühlten, der gleich dem eines Gemenges von KCl und Na Cl-Kristallen ist. Mit der Zeit nimmt der Wärmeinhalt der abgeschreckten Schmelzen ab und um so schneller, bei je höherer Temperatur sie aufbewahrt werden.

Aus ihren gemischten Lösungen in Wasser scheiden sich KCl und Na Cl in fast reinem Zustande aus[3].

---

1) Kurnakow und Zemczuznyj, Zeitschr. f. anorg. Chemie 52, 186, 1907.
2) Nacken, Sitzungsb. d. Akademie in Berlin 1918, Februar-Heft S. 192.
3) Knickmeyer, Zeitschr. f. phys. Chemie 21, 62, 1896.

KCl und Na Cl bilden also bei höheren Temperaturen zwischen 660° und 500° eine stabile lückenlose Mischkristallreihe, und da weder KCl noch Na Cl einen Umwandlungspunkt besitzt, so können die beschriebenen Vorgänge bei der Abkühlung ihrer Mischkristalle nur als eine Entmischung derselben gedeutet werden.

Bei flüssigen Mischungen ist der Grund einer Entmischung bei sinkender Temperatur darin zu suchen, daß die Anziehungen der Moleküle je einer der beiden Komponenten die beider Molekülarten überwiegen, wodurch die teilweise Trennung beider Komponenten bewirkt wird. Dasselbe würde auch für anisotrope Mischungen zutreffen können. Dann sollten mit sinkender Temperatur in den Mischkristallen zwischen den Molekülen des KCl oder denen des Na Cl die Attraktionskräfte wachsen oder die zwischen den KCl- und Na Cl-Molekülen wirkenden Kräfte stark abnehmen.

Die an KCl- und Na Cl-Kristallen beobachteten Ätzfiguren weisen darauf hin, daß bei gewöhnlicher Temperatur die den KCl-Kristall zusammenhaltenden Kräfte sich von dem des Na Cl-Kristalls wesentlich unterscheiden, oder in bestimmterer Ausdrucksweise, daß im KCl-Kristall gewisse Gitterbezirke miteinander fester zusammenhängen, während im Na Cl solche Unterschiede nicht bestehen. Man könnte auch sagen, daß das Molekulargewicht des KCl ein Multiplum des normalen ist, während das des Na Cl mit dem normalen übereinstimmt, wenn man nicht die Bezeichnung Molekül für diejenigen Atomgruppen reservieren will, die bei der Diffusion im Gitter sich geordnet bewegen. Wenn im KCl diese Gitterbezirke bei hohen Temperaturen nicht bestehen, sondern erst bei tieferen Temperaturen sich bilden, so würde in ihrer Entstehung der Grund für die Entmischung zu suchen sein.

Über die Ätzfiguren an Na Cl hat R. Brauns[1]) folgende Angaben gemacht. In feuchter Luft bilden sich auf den Würfelebenen des Steinsalzes quadratische Ätzgrübchen. In der Regel verlaufen die Umrisse dieser Ätzgrübchen parallel den Würfelkanten Fig. 1, nur bei einem Kristall wurde eine Lage parallel den Diagonalen der Würfelebene Fig. 2 beobachtet.

Da die das Ätzgrübchen begrenzenden Flächen nicht bei derselben Konzentration der einwirkenden Lösung entstanden sind, so ist ihre Neigung zur Würfelebene eine Funktion dieser Konzentration und daher im allgemeinen eine ziemlich unbestimmte.

---

1) R. Brauns, Neues Jahrb. f. Min. 1886 I, S. 244, 1889 I, S. 113.

1) G. Linck, Tschermacks min. petr. Mitteilungen 12, 1891, S. 82.

Die Untersuchung der Lichtfiguren der Ätzgrübchen lehrt, daß ihre Flächen den Umrissen der Ätzfigur nicht entsprechen und in der Regel Flächen doppelter Krümmung sind, die von Ebenen mit hohen Indices unterbrochen sind. Nur eine Ätzfigur, bei deren

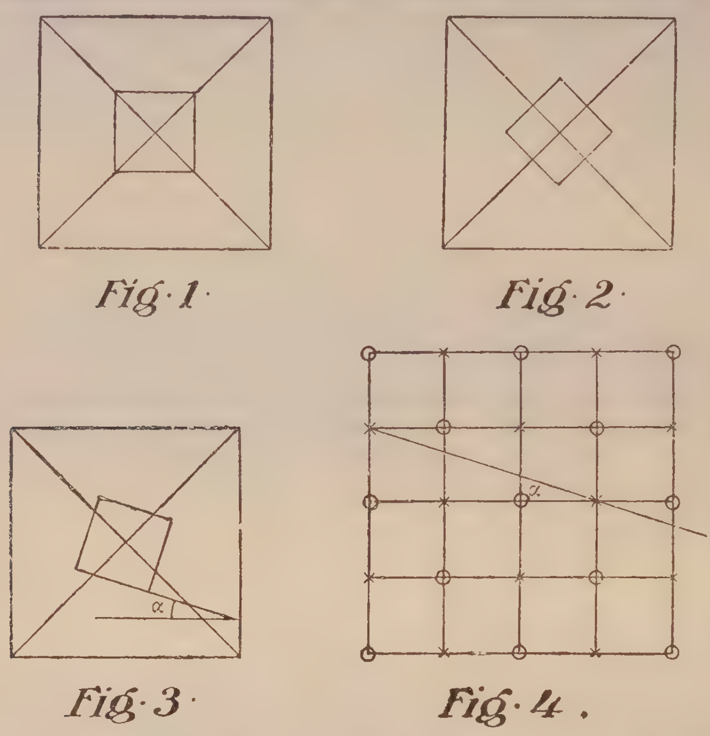


Fig. 1. Fig. 2. Fig. 3. Fig. 4.

Entstehung die Konzentration des Ätzmittels an allen Punkten der Oberfläche dieselbe ist, eine Ätzfigur konstanter Konzentration, würde von Ebenen begrenzt sein, welche der Bedingung genügen, daß in gleichwertigen Richtungen die Lösungsgeschwindigkeit dieselbe ist. Wenn auch die Flächen der Ätzgrübchen nicht der Bedingung konstanter Konzentration entsprechen, so trifft diese doch für die Schnittlinien der Ätzgrübchenflächen mit der geätzten Ebene sehr angenähert zu. Man wird also diese Schnittlinien, die Umrisse der Ätzfigur, auf der geätzten Kristallebene als eine Figur, die konstanter Konzentration des Ätzmittels entspricht, betrachten dürfen.

Die ebenfalls bei der Einwirkung von feuchter Luft auf KCl entstehenden Ätzfiguren unterscheiden sich von denen des NaCl nach Brauns[1]) sehr wesentlich und G. Linck[2]) hat diese Angaben

1) l. c.

2) G. Linck, Tschermacks min. petr. Mitteil. 12, 1891, S. 82.

bestätigt. Die Umrisse der Ätzfiguren bilden mit der Würfelkante nach Brauns die Winkel α Fig. 3 von 11 bis 18°, während Linck bei zwei Bestimmungen den Wert von α zu 17° fand.

Nach diesen Feststellungen betreffend die Ätzfiguren des Na Cl und KCl war es recht überraschend, daß W. L. Bragg[1]) für beide dieselbe Gitterart fand, nämlich das einfachste kubische Gitter, dessen Elementarwürfel an den Ecken Cl- und K- oder Na-Atome trägt. Die Kantenlänge dieses Würfels ist beim KCl $6{,}27.10^{-8}$ und beim Na Cl $5{,}49.10^{-8}$ cm. Die Erwartung, daß dem KCl ein plagiedrisch-hemiedrisches Gitter zukommt, ist also nicht erfüllt. Beiden kommt dieselbe Gitterart holoedrischer Symmetrie zu.

Sucht man die Umrisse einer Ätzfigur auf Grund der atomistischen Gitterstruktur zu verstehen, so liegen folgende Möglichkeiten vor. Auf einer Gitternetzebene der Würfelebenen ist die Verteilung der Cl- und Na-Atome die in Fig. 4 angedeutete. In den Richtungen parallel den Würfelkanten folgen sich abwechselnd die durch Kreuze und Kreise gekennzeichneten beiden Atomarten. Wenn das Ätzmittel ein Atompaar, Na Cl, dem Gitterverbande entreißt, so werden dadurch die benachbarten Atompaare besonders reaktionsfähig. Je nachdem wie die Stellen der Atome, die besonders reaktionsfähig geworden sind, ihre Empfindlichkeitsbezirke, orientiert sind, pflanzt sich der Angriff auf der Oberfläche fort. Wenn ihre Orientierung parallel den Würfelkanten ist, so entstehen Furchen parallel den Würfelkanten, die sich zu Ätzfiguren mit Umrissen parallel den Würfelkanten gestalten werden, wenn die Empfindlichkeitsbezirke parallel den Diagonalen der Würfelebenen orientiert sind, so werden Ätzfiguren mit Umrissen parallel diesen Diagonalen entstehen. Die Orientierung der Empfindlichkeitsbezirke kann sich mit der Natur des Ätzmittels und durch Gegenwart fremder Stoffe leicht ändern. Hierauf dürfte wohl der Befund von R. Brauns zurückzuführen sein, daß bei einem Na Cl-Kristall die Ätzgrübchen sich in der Diagonalstellung bei allen andern ausnahmslos in Stellung parallel den Würfelkanten befanden. Der Einfluß des Ätzmittels ist besonders von F. Becke[2]) eingehend verfolgt worden. Beim Bleiglanz (PbS), dem ebenfalls das einfachste reguläre Gitter zukommt, können Ätzfiguren entweder der Würfelkanten- oder der Diagonalstellungen entstehen. Bei der Einwirkung kalter konzentrierter Salzsäure entstehen die der ersten, bei der Einwirkung 20 %iger Säure oder von Kalilauge die der zweiten Orientierung.

1) W. L. Bragg, Zeitschr. f. anorg. Chemie 90, 1915, S. 246.
2) F. Becke, Tschermacks min. petr. Mitteil. 6, 1885, S. 237.

Zur Deutung anders orientierter Ätzfiguren, wie sie beim KCl auftreten, muß außer der Orientierung der Empfindlichkeitsbezirke nach den Würfelkanten noch eine besondere Eigentümlichkeit des Auflösungsvorganges angenommen werden. Aus der Netzebene der Würfelebene werden vier Atompaare, die auf zwei benachbarten Geraden parallel den Würfelkanten liegen, vom Agens aus dem Gitterverbande genommen und dann wieder vier Atompaare, die aber nicht auf demselben Geradenpaar, sondern auf einer dieser beiden Geraden und der ihr benachbarten liegen. Wenn die Entfernung der Atompaare aus dem Gitterverbande in dieser Weise sich fortsetzt, so würde eine gezackte Furche entstehen. Die Richtung dieser Furche macht mit der Würfelkante den Winkel α. Wenn die Kantenlänge des Elementarwürfels a beträgt, so würde, wenn nach Entfernung von je vier Atompaaren die beschriebene Verschiebung der folgenden Atompaare eintritt,

$$\operatorname{tang} \alpha = \frac{a}{4a} = \frac{1}{4}$$

sein; also $\alpha = 14^0 2'$. Würde die Verschiebung nach der Entfernung von je drei Atompaaren stattfinden, so wäre

$$\operatorname{tang} \alpha = \frac{a}{3a} = \frac{1}{3};$$

also $\alpha = 18^0 26'$. R. Brauns fand α zwischen 11 bis 18^0 schwankend und G. Linck fand $\alpha = 17^0$.

Der Grund jener Verschiebung ist vielleicht in einem festen Zusammenhang von je vier Atompaaren zu suchen. Der festere Zusammenhang von je drei Atompaaren ist unwahrscheinlicher. Was für die Reihen parallel einer Würfelkante gilt, muß auch für die Reihen parallel den drei zueinander senkrechten gelten. Man kommt also zum Resultat, daß im Gitter des KCl Würfel von $(4a)^3$ oder $(3a)^3$ existieren, deren Atome nicht nur durch die von Atom zu Atom wirkenden Kräfte zusammengehalten werden, sondern noch durch andere Kräfte, die die Atompaare jener Bezirke zusammen halten.

Die Bestimmung dieser Bezirke ist keine eindeutige. Da

$$\operatorname{tang} \alpha = \frac{a}{4a} = \frac{2a}{8a} = \frac{4a}{16a} \text{ usw.}$$

ist, so könnten die Bezirke statt der Größe $(4a)^3$ auch die von $(8a)^3$ oder $(16a)^3$ usw. haben. Wenn mit steigender Temperatur diese Bezirke im KCl sich gewissermaßen auflösen, so wird das KCl seinem atomistischen Aufbau nach dem NaCl, in dem jene

Bezirke auch bei tieferen Temperaturen fehlen, immer ähnlicher. Dem entsprechend sollten die Ätzfiguren des KCl über 400° wie die des Na Cl orientiert sein.

Wenn in KCl mit sinkender Temperatur die Bildung jener Bezirke in einem weiten Temperaturintervall vor sich geht, dagegen in den Mischkristallen in einem viel kleineren, so würde im KCl selbst die Wärmeentwicklung bei der Bildung jener Bezirke nicht merklich werden, während sie in den Mischkristallen merklich wäre und noch durch die Entmischungswärme vergrößert würde.

Lückenlose Mischkristallreihen, die sich mit recht verschiedenen Geschwindigkeiten entmischen, bilden auch andere Kalium- und Natriumsalze bei der Kristallisation aus ihren Schmelzen, wie beim KBr und Na Br[1]), KJ und Na J[1]) sowie KBO_2 und Na BO_2[2]). Dementsprechend ist zu erwarten, daß sich auch bei diesen Salzpaaren die beim KCl und Na Cl bestehenden Unterschiede in den Ätzfiguren finden werden.

---

1) Kurnakow, Zemczuzny u. Rambach, Z. f. anorg. Chemie 65, 1909, S. 403.
2) van Klooster, Zeitschr. f. anorg. Chemie 69, 1911, S. 131.

Versuche über die Erblichkeit des in farbigem Lichte erworbenen Farbkleides der Puppen von Pieris brassicae.

Zweite vorläufige Mitteilung.

Von

Bernhard Dürken, Göttingen.

Vorgelegt von Herrn Ehlers in der Sitzung vom 31. Oktober 1919.

In diesen Nachrichten habe ich vor etwa einem Jahr eine vorläufige Mitteilung veröffentlicht über die Wirkung farbigen Lichtes auf Puppe und Falter von *Pieris brassicae* und die Beschaffenheit der unbeeinflußten Nachkommen[1]. Die Versuche sind inzwischen fortgesetzt worden und haben auch bereits weitere Ergebnisse geliefert. Zu einer einigermaßen abschließenden Beurteilung dieser Ergebnisse sind aber noch besondere Versuchsreihen erforderlich, die, wenn auch zum Teile bereits eingeleitet, erst in längerer Zeit zu Ende geführt werden können.

Die Fortpflanzung von Tagschmetterlingen in der Gefangenschaft ist mit verhältnismäßig großen Schwierigkeiten verbunden, so daß bereits mehr als eine Versuchsreihe zu wiederholtem Male von vorn begonnen werden mußte. Insbesondere hat auch die Ungunst der Witterungsverhältnisse, welche in den letzten Sommern gerade während der Fortpflanzungsperioden herrschte, verheerend in meinen Zuchten gewirkt. Aus diesen und anderen Umständen heraus mag es begründet erscheinen, das Endergebnis aller Versuche nicht abzuwarten, sondern in einer zweiten vorläufigen Mitteilung das wichtigste Ergebnis eines weiteren Teilversuchs zu veröffentlichen. Auch die ausführliche Schilderung

---

1) Nachr. d. K. Gesellsch. d. Wiss. zu Göttingen, Mathem.-physikal. Kl., 1918.

der in der ersten Mitteilung behandelten Versuche soll zunächst noch zurückgestellt und im Zusammenhang mit den laufenden Versuchen später erfolgen, vor allem auch, weil dadurch manche Wiederholungen vermieden werden können.

Das für die folgende Darlegung hauptsächlich in Betracht kommende Ergebnis jener ersten vorläufigen Mitteilung ist ganz kurz gefaßt folgendes:

Wachsen Raupen des Kohlweißlings (*Pieris brassicae*) in orangefarbenem Lichte auf und erfolgt die Verpuppung unter den gleichen Bedingungen, so wird im Integument der Puppe das schwarze und weiße Pigment außerordentlich stark reduziert, so daß die meisten der Puppen infolge Durchscheinens der Körpersäfte und -gewebe grün oder doch — wenn noch etwas weißes Pigment vorhanden ist — stark grünlich erscheinen mit nur kleinen schwarzen Zeichnungselementen. In Zuchten aus nichtfarbiger Umgebung beträgt der Anteil dieser Puppen (Gruppe *B*) an der Gesamtzahl nach dem Durchschnitt einer größeren Anzahl von Zuchten rund 7 v. H., derjenige der Puppen mit viel Weiß und mindestens ziemlich viel Schwarz (Gruppe *A*) 93 v. H. In einer besonderen Kontrollzucht, bei der die Verpuppung auf dem gleichen grauen Untergrund erfolgte wie bei der unten erwähnten Hellzucht der P_2-Generation, betrug allerdings der Anteil der Gruppe *B* rund 8,8 v. H. Der Unterschied vom Durchschnitt ist aber, wie man sieht, nur gering; in anderen Kontrollzuchten traten weniger als 7 v. H. in der Gruppe *B* auf.

Für die genannten Versuchszuchten in orangenem Licht lauten diese Zahlen: $A = 22{,}2\,\%$; $B = 77{,}8\,\%$.

Schlüpfen nun die Falter der Gruppe *B* ebenfalls in orangenem Licht, so ergeben deren Nachkommen an Puppen der Gruppe *B*:

unter Fortfall der Versuchsbedingungen: $B = 42{,}8\,\%$;

unter Fortdauer der Versuchsbedingungen: $B = 94{,}8\,\%$.

Hier handelt es sich also um die grundsätzlich und praktisch äußerst wichtige Frage nach der Vererbung des Erworbenen. Das mitgeteilte Ergebnis genügte allerdings noch nicht, um über das Wie und Wodurch dieser Übertragung auf die Nachkommen eine Entscheidung zu ermöglichen. Zur schärferen Erfassung des Problems wurden nun u. a. zwei Versuchsreihen begonnen, welche bezweckten, den Zeitpunkt der wirksamen Einwirkung auf die Eltern näher zu bestimmen. Denn von dieser Bestimmung hängt vor allem die Beantwortung der Frage nach dem Wie der Übertragung ab. Mehr über die Tragweite dieser Bestimmung soll hier zunächst nicht ausgeführt werden.

Nun ist schon länger bekannt, daß zur Erzielung der grünen Puppenfärbung nicht die ganze Aufzucht der Raupen in orangenem Licht zu erfolgen braucht, sondern daß es genügt, die Raupen kurz vor der Verpuppung in solches Licht zu übertragen. Die Frage war also: wird auch diese Grünfärbung der so erzielten Gruppe *B* auf die Nachkommen übertragen, insbesondere auch dann, wenn gleich nach Vollendung der Puppenfärbung die Wirkung des Versuchsfaktors ausgeschaltet wird und die Entwicklung des Falters ohne denselben erfolgt?

Ferner: wenn die Grünfärbung der Puppen auf die Nachkommen übertragen wird, ist dann nicht auch die weiß-schwarze Normalfärbung erblich?

Um das zu entscheiden und zugleich die oben erwähnte Zeitbestimmung vorzunehmen wurden zwei Versuchsreihen angesetzt:

1) Grüne Puppen (= P_1-Generation), welche lediglich während der eigentlichen Verpuppungszeit in orangenem Licht waren, wurden in nichtfarbiger Umgebung überwintert und die Nachkommen der daraus bei Sonnenlicht geschlüpften Falter aufgezogen (= P_2-Generation).

2) Puppen mit viel Schwarz und Weiß aus nichtfarbiger Umgebung wurden gleich nach der Ausfärbung in orangenes Licht gebracht, worin sie überwinterten. Auch die Falter schlüpften in diesem Licht. Die Aufzucht der Nachkommen liefert den notwendigen Gegenversuch zu der ersten Versuchsreihe.

Hier soll nur von dem Ergebnis dieser ersten Reihe die Rede sein; die zweite Versuchsreihe mußte im letzten Sommer (1919) zum dritten Male begonnen werden; ein brauchbares Ergebnis ist frühestens im Sommer 1920 zu erhoffen.

I. Beschaffenheit der P_1-Generation[1]).

Die Raupen der P_1-Generation wuchsen in nichtfarbiger Umgebung auf, wenn man von der grünen Färbung des Futters absieht. Sie befanden sich in einem Käfig aus Drahtgewebe, so daß diffuses Tageslicht allenthalben Zutritt hatte; der Käfig war innen dunkelgrau gestrichen. Zur Verpuppung kamen die reifen Raupen in die gleichen mit orangefarbenen Scheiben versehenen Zuchtkästen, wie sie in den Vorjahren benutzt waren. Alsbald nach der Ausfärbung wurden die Puppen abgenommen und in die beiden schon erwähnten Gruppen *A* und *B* gesondert. Die Gruppe *B*

1) Belegexemplare der im folgenden erwähnten Versuchsreihen befinden sich im Phyletischen Museum zu Jena.

wurde in den leeren Drahtgewebekäfig zurückgebracht und in einem Zimmer, dessen Fensterläden meist geschlossen waren, überwintert. Das Zimmer war nicht völlig dunkel, sondern durch eine offene Tür und durch Spalten drang stets so viel Licht ein, daß nur eine starke Dämmerung herrschte; auch wurden die Läden öfter vorübergehend geöffnet.

Im Frühjahr wurde der Käfig mit den Puppen ins Freie gebracht, so daß die schließliche Ausbildung der Falter und ihr Schlüpfen bei vollem Tageslicht, zum Teil in direktem Sonnenlicht erfolgte. Auch die Fortpflanzung und die Eiablage vollzogen sich im Freien bei ungehindertem Lichtzutritt. Die aus diesen Eiern gezogenen Raupen bezw. Puppen stellen die P_2-Generation der folgenden Versuche dar.

Die P_1-Generation ergab vorwiegend Puppen der Gruppe B, und zwar:

A =	59	
B =	100	$B = 62{,}8\,\%$
Gesamtzahl	159	

II. Beschaffenheit der P_2-Generation.

Aus den Nachkommen der Gruppe B der P_1-Generation wurden drei Zuchten angelegt. Für alle drei wuchsen die Raupen unter gleichen Bedingungen heran, wie diese Bedingungen auch übereinstimmten mit denen der P_1-Zuchten. Die Verpuppung erfolgte unter dreierlei verschiedenen Bedingungen:

1) in diffusem weißen Tageslicht auf grauem Untergrund;
2) in verdunkelten, schwarz ausgekleideten Behältern;
3) in demselben orangenen Licht und in denselben Zuchtkästen wie bei der Eltern-Generation.

Erst die reifen Raupen wurden in die Verpuppungsbehälter gebracht. Versuche Nr. 1 und 2 kann man als Zuchten mit Fortfall der Versuchsbedingungen zusammenfassen, da in ihnen die Raupen nicht in dem spezifischen Sinne der Eltern-Generation beeinflußt wurden, während in Zucht Nr. 3 die Versuchsbedingungen bestehen blieben.

1. Versuche mit Fortfall der Versuchsbedingungen.

A =	74	
B =	63	$B = 45{,}9\,\%$
Gesamtzahl	137	

Bei der Würdigung dieser Zahlen ist zu berücksichtigen, daß in den Hellzuchten ein, wenn auch nur schwacher Einfluß auf die Puppenfärbung von den grünen Blättern der Futterpflanzen ausgeht, der in ähnlichem Sinne wirkt wie orangenes Licht. Besonders trifft das zu für diejenigen Raupen, welche sich unmittelbar am Boden des Behälters zwischen den Kohlblättern verpuppt haben. Für diejenigen, welche sich an den hellgrauen Seitenwänden des Behälters angeheftet haben, ist der Einfluß der am Boden liegenden grünen Blätter erfahrungsgemäß sehr gering, da die Beschaffenheit des Anheftungsgrundes für die Puppenfärbung maßgebend ist. Andererseits muß man aber bedenken, daß in den Dunkelzuchten vielleicht zum Teil ein gegenteiliger Einfluß auf die Puppenfärbung ausgeübt wurde. Da bei ihnen kein vollständiger Lichtabschluß vorlag, sondern durch Spalten am Deckel und durch Lüftungsgitter immerhin etwas Licht eindringen konnte, so dürfte hier wenigstens bei einem Teil der Puppen die schwarze Umgebung zur Wirkung gekommen sein, wodurch erfahrungsgemäß die Bildung schwarzen Pigments begünstigt wird.

Somit dürfte es richtig sein, wie es oben geschehen ist, sämtliche Puppen der Hell- und Dunkelzuchten zur Beurteilung des Erfolges heranzuziehen, da sich die gegensätzlichen Einflüsse hier und dort die Wage halten dürften.

Berücksichtigt man nur diejenigen Puppen der Hellzucht, welche von den grauen Wänden des Verpuppungsbehälters stammen, die also sicher in keiner Weise spezifisch beeinflußt sind, so lauten die in Betracht kommenden Zahlen:

$A = 30$
$B = 28$ $\qquad B = 48{,}2\,\%$
Gesamtzahl 58

Man sieht, daß dann das Ergebnis zu Gunsten der Gruppe B verschoben ist. Weitere Zahlen sollen hier zunächst nicht mitgeteilt werden.

2. Versuche mit Wiederholung der Versuchsbedingungen.

Die Verpuppung der P_2-Generation erfolgte unter denselben Bedingungen wie die ihrer Eltern.

$A = 1$
$B = 56$ $\qquad B = 98{,}2\,\%$
Gesamtzahl 57

Die Zusammenstellung der mitgeteilten Zahlen ergibt für den Anteil der Gruppe B an den einzelnen Zuchten:

1) Kontrollzuchten durchschnittlich . . . . . $B = 7{,}0\,\%$
2) einmalige Wirkung von orangenem Licht bei der Verpuppung (P_1) . . . . $B = 62{,}8\,\%$
3) wiederholte Wirkung von orangenem Licht bei der Verpuppung der P_2-Generation . . $B = 98{,}2\,\%$
4) Unbeeinflußte Aufzucht der P_2-Generation durchschnittlich . . . . $B = 45{,}9\,\%$

Damit ist also mit anderen Worten erwiesen, daß die Puppenfärbung in hohem Maße von der Färbung der Eltern abhängt. Insbesondere wird die durch Verpuppung in orangenem Licht erworbene Grünfärbung in einem beträchtlichen Grade auf die unbeeinflußten Nachkommen der ersten Generation übertragen. Wiederholt sich aber bei der Nachkommen-Generation der Einfluß des abändernden Faktors, so tritt eine erhebliche Steigerung der Wirkung ein; man kann sagen, daß dann so gut wie alle Nachkommen die aufgedrungene Beschaffenheit angenommen haben. Wenn bei Fortfall der Versuchsbedingungen ein erheblicher Anteil der P_2-Generation nicht den direkten Vorfahren (P_1-Generation) gleicht, sondern in der Richtung der ursprünglichen Vorfahren variiert (Gruppe A), so zeigt das lediglich, daß durch einmalige Einwirkung des abändernden Faktors die ursprünglichen Vererbungstendenzen, um es hier einmal so zu nennen, noch nicht völlig überwunden sind.

Wichtig an den jetzt mitgeteilten Ergebnissen und neu gegenüber den früheren Versuchen ist vor allem, daß eine ganz vorübergehende Einwirkung genügt, um eine übertragbare Abänderung zu erzeugen; ferner daß diese Einwirkung nur auf ein Stadium stattgefunden hat, welches nicht nur zeitlich, sondern auch qualitativ-entwicklungsgeschichtlich entfernt liegt von der endgültigen Ausbildung der Geschlechtsprodukte dadurch, daß die „Reifung" des Falters und die Fortpflanzung nach Aufhören der Versuchsbedingungen vor sich gegangen sind.

Es liegt nun außerordentlich nahe, daraus zu folgern, daß nicht die Fortpflanzungszellen selbst unmittelbar auf den abändernden Faktor reagiert haben, sondern daß sie durch die ihrer endgültigen Ausbildung voraufgegangene Abänderung des umgebenden Mediums erst nachträglich selbst jene Beschaffenheit erlangt haben, welche für die Übertragung der Abänderung notwendig ist. Um das endgültig entscheiden zu können, muß das Ergebnis weiterer Versuche abgewartet werden.

Immerhin darf man auch jetzt schon mit großer Wahrscheinlichkeit eine Vermutung aussprechen über das Mittel, durch welches die erworbene Abänderung auf die Nachkommen übertragen wird. Die Verpuppung in orangenem Licht verursacht eine spezifische chemische Beschaffenheit des Integuments, die sich in der besonderen Art des Verhaltens der Pigmente äußert, und die gleiche chemische Beschaffenheit der Körpersäfte. Auf irgend einem Wege dürften auch die heranwachsenden Keimzellen diese Beschaffenheit erwerben. Wenn dann dadurch die elterngleiche Ausgestaltung der Nachkommen bedingt wird, so ist es klar, daß die Übertragung des Erworbenen hier nicht durch die Modifikation besonderer, etwa im Gametenkern anzunehmender Vererbungsträger bewerkstelligt wird, sondern durch die chemische Allgemeinbeschaffenheit der Gameten. Wir würden es also hier nicht mit karyogener (vom Gametenkern ausgehender) Vererbung zu tun haben, sondern, wie ich derartiges an anderem Orte genannt habe, mit plasmogener. Jedoch auch hierüber müssen neue Versuche weiteren Aufschluß bringen.

Über die Ausbildung des Oberdevon in der Umrandung des Siegerländer Blockes.

Von

R. Wedekind-Marburg.

Vorgelegt von Herrn Stille in der Sitzung vom 31. Oktober 1919.

1. Einführung.

Ein Schichtenschnitt durch das Oberdevon hat im nordöstlichen Sauerlande folgenge Zusammensetzung (v. o. n. u.):

Padberger Schiefer

— — — — — —

Fossley

— — — — — —

Cephalopodenkalke.

— — — — — — — —

Unterlage: Riff- oder Flinzplatte.

Durch die biostratigraphische Gliederung der Cephalopodenkalke wurden bestimmte Zeitspannen genauer festgelegt und damit ein Mittel gewonnen, die räumliche Entwicklung der oberdevonischen Sedimente zu verfolgen. Da sich am Nordrande des rheinischen Gebirges von Marsberg bis nach Balve im Fossley und den Padberger Schiefern oder den diesen entsprechenden Wocklumer Kalken von denen der Cephalopodenkalke abweichende Fossilien fanden, konnten Fossley und Padberger Schiefer (= Wocklumer Kalk) nicht vom Alter der Cephalopodenkalke sondern nur jünger sein. Damit war ich zu einer Anschauung gekommen, die sich abgesehen von mancherlei Modifikationen an die Ausführungen Denckmanns anschließt.

Dieser Anschauung der Altersverschiedenheit dieser Bildungen stehen nach wie vor ältere und neuere Beobachtungen gegenüber, die nunmehr von der gewonnenen sicheren Grundlage aus zu überprüfen sind.

2. Der Sedimentationsraum.

Für die Betrachtung fossiler Sedimente, also auch des Oberdevon, ist die Kenntnis des Sedimentationsraumes nebst den dazu gehörigen Denudationsgebieten von grundlegender Bedeutung. Die Möglichkeit für die Bestimmung des Sedimentationsraumes der oberdevonischen Sedimente ist durch die besonderen Verhältnisse des oberen Mitteldevon gegeben. Dieses besteht aus dunklen, meist krystallinen, häufig Tentaculiten führenden Plattenkalken, die mit Tonschiefern wechsellagern. Die Gesamtheit dieser Sedimente wird in manchen Gebieten als Flinz bezeichnet, ein Ausdruck, den ich in erweitertem Sinne übernehme. Die zentralen Flinzkalke gehen in bestimmten Richtungen in Riffkalke als Randbildungen über, welche sich zu geschlossenen Saumriffen zusammenschließen, zu denen noch Wallriffe und Atolle hinzutreten. Im rheinischen Meeresbecken sind zwei Saumriffe zu unterscheiden, von denen das eine die Südküste des nordatlantischen Kontinentes begleitet. Der allgemeine Verlauf dieses Riffes ist durch folgende Punkte gegeben: Barmen, Iserlohn, Balve, Warstein und Brilon. Das Südriff hat eine ganz andere Verbreitung. Es zieht in fast geschlossenem Zuge von Weilburg über Wetzlar nach Giessen, umgibt das Ostende des Taunus und ist dann auch im Süden des Taunus in sehr vereinzeltem Vorkommen bekannt geworden. Das nördliche Verbreitungsgebiet ist ebenfalls nur durch vereinzelte Vorkommen gekennzeichnet, die aber zeigen, daß das Saumriff den nördlichen Teil der Dillmulde durchzieht und hier den Südostrand des Siegerländer Blockes begleitet. In der östlichen Umrandung des Siegerländer Blockes sind Riffkalke und Bildungen des oberen Mitteldevons bisher nicht bekannt geworden, Erst in der Nordumrandung treten Riffkalke in der Attendorner Doppelmulde wieder hervor. Durch das Vorkommen von Paffrath wird der Anschluß mit der Eifler Dolomitbucht gewonnen.

Diese Riffe zeigen mit voller Deutlichkeit die Südküste des nordatlantischen Kontinentes und wenigstens in großen Zügen den Verlauf der Ostküste der mitteleuropäischen Insel an. Im heutigen Gebiete der Dill und Lahn sprang gegen das Innere der Insel die hessische Bucht und im Süden des Taunus die Taunusbucht vor.

3. Die oberdevonischen Sedimente in der Umrahmung des Siegerländer Blockes.

Da die Bildung klastischer Sedimente unmittelbar abhängig ist von der Abtragung der Gebiete, die das Material für die Sedimentation liefern — Sedimentation ist eine Funktion der Denudation — sollen hier nicht beliebige, sondern ihrer palaeogeographischen Position nach bekannte Schichtenschnitte zunächst betrachtet werden. Dabei wird von Schnitten ausgegangen, die sich in der Umrandung des Siegerländer Blockes finden.

a) Die Dillmulde.

In der Dillmulde tritt ein Komplex roter Schiefer und zwischen diesen und dem Unterkarbon eine Folge mächtiger Sandsteine auf. Dazu gesellen sich im Hangenden des Mitteldevon Kramenzelkalke, die infolge ihres selten fehlenden Fossilinhaltes leicht als oberdevonisch zu erkennen sind. Infolge von Überschiebungen und Durchschiebungen, dadurch hervorgerufen, daß Kalk und Schiefer sich dem tangentialen Druck gegenüber verschieden verhalten haben, gibt es kaum zusammenhängende Profile. Jedes dynamisch selbständige Schichtenglied besitzt auch ein eigenes tektonisches Gepräge! Dieser Sachverhalt wird noch dadurch erschwert, daß die Cephalopodenkalke nach S in eine andere Facies übergehen, die Annulatakalke der Dillmulde finden sich in der Lahnmulde als Annulataschiefer, und daß die Cephalopodenkalke vielerorts außerdem fehlen.

Nach meinen im nördlichen Sauerlande gesammelten Erfahrungen mußte die Schichtenfolge der Dillmulde folgendermaßen gedeutet werden:

a. Obere Sandsteine (= Pönsandsteine)
b. Rotschiefer (= Fossley)
c. Cephalopodenkalke
Mitteldevonische Riff- oder Flinzkalkplatte.

Da in der näheren Umgebung von Bicken Cephalopodenkalke einerseits und Rotschiefer nebst den oberen Sandsteinen anderseits in einem engbegrenzten Gebiete zusammen auftreten und infolgedessen auch altersverschieden sein müssen, konnte bei Bicken das Alter der Rotschiefer bestimmt werden.

Im Sommer 1918 gelang es mir, in dem bekannten Steinbruche an der Bickener Landstraße einen Ausgangspunkt zu gewinnen. Außer den Kalken der Mauticocerasstufe treten hier jüngere oberdevonische Sedimente auf, die zuunterst aus festen Kalken bestehen

und gegen das Hangende in einen sehr tonreichen Kramenzelschiefer übergehen. Die hangenden Kramenzelschiefer führen eine reiche Fauna, die aus Clymenien der Gruppe der Clymenia laevigata und Gonioclymenien besteht. Die Fauna läßt bei der vollständigen Übereinstimmung mit dem Balver Vorkommen keinen Zweifel darüber, daß die Kramenzelkalke der Laevigatastufe entsprechen.

Diese Altersbestimmung wurde weiterhin noch dadurch bestätigt, daß es gelang, in dem Liegenden der Kramenzelschiefer eine kleine Fauna zu sammeln, die aus kleinen, aber selten gut erhaltenen Platyclymenien besteht; wieder handelt es sich um die gleichen Arten wie im Balver Oberdevongebiet [1]).

Zur Ergänzung dieses Befundes wurden weitere Fundpunkte und insbesondere die reichen Materialien des Marburger Institutes herangezogen. Die liegenden Teile der oberdevonischen Kalke sind besonders durch die Untersuchungen von Lotz und die jahrelangen Aufsammlungen E. Kaysers bekannt geworden, ohne daß der stratigraphische Wert dieser Aufsammlungen bisher gewürdigt wäre.

Die Roteisensteine an der Basis des Oberdevon zeigen die reiche Entfaltung von Pharciceras und Gephyroceras. Es handelt sich also um die Zone I α. Ob sich in dieser Zone noch Vertreter der Gattung Maeneceras finden, konnte bisher nicht festgestellt werden. Da Maeneceras und Manticoceras genetisch keine Beziehungen zu einander zeigen, ist das Vorkommen von Nachläufern der Gattung Maeneceras immerhin möglich. Die Zugehörigkeit der Pharciceraszone zum Oberdevon wird aber dadurch nicht berührt, da das Oberdevon mit dem Einsetzen der Manticoceratiden beginnt. Im Hangenden folgen plattige Kalke (Sessacker) von roter Farbe, die die Adorfer Fauna der Zone I γ und häufig Manticoceras cordatum führen. Bei Bicken setzt dann eine Folge von dunklen Schiefern und Kalken mit dunklen fossilführenden Kalkknollen ein mit den Zonenfossilien und der Fauna der Zone I δ.

Zwischen dieser Zone und den Postprolobitesschichten sind die Cheiloceras- und Prolobitesschichten in der Dillmulde petrographisch und faunistisch normal entwickelt. Auch die Grenzschichten mit Pseudoclymenia Sandbergeri, die an der Unterkante der Prolobitesstufe liegen, konnte wenigstens einstweilen nach den vorhandenen Aufsammlungen vom Sessacker festgestellt werden.

Es ist demnach der Nachweis erbracht worden, daß in der

1) Ausführliche Fossillisten wird die Dissertation meines Assistenten O. H. Schindewolf bringen.

Dillmulde die gleiche Schichtenfolge vorhanden ist, wie am Nordrande des rheinischen Gebirges, und daß außerdem die Rotschiefer stratigraphisch die gleiche Lage einnehmen wie am Nordrande des Rheinischen Gebirges uud als Fossley zu bezeichnen sind. Es kann bei Bicken also nicht die Rede davon sein, daß Rotschiefer und Fossley die Cephalopodenkalke vertreten.

Zwischen Haiger und Donsbach sind die Verhältnisse modifiziert. Hier treten in der Unterlage der Rotschiefer Flinzkalke hervor, die ein jungmitteldevonisches resp. altoberdevonisches Alter ($1\alpha, \beta$) haben. Cephalopodenkalke fehlen. Es gelang hier, ein Konglomerat an der Basis des Fossley aufzusuchen, daß aus gerundeten Brocken stark zersetzter oberdevonischer Kalke besteht. Mithin sind hier zweifellos oberdevonische Cephalopodenkalke abgetragen. Der Fossley transgrediert hier in ausgesprochener Weise.

b) *Das Rimberggebiet nördlich von Marburg.*

Von ganz besonderem Interesse ist der östliche Teil des von H. Lieber bearbeiteten Rimberggebietes. H. Lieber hat seine Untersuchungen nicht zum Abschluß bringen können. Daher ist es auch erklärlich, daß er mancherlei übersehen oder irrtümlich gedeutet hat. Er erkannte zunächst folgende Schichtenfolge:

Culm
Pönsandstein
Buchenauer Schichten
Schiefer des oberen Mitteldevon
Wissenbacher Schiefer.

Stratigraphisch deutete er diesen Schichtenschnitt, in dem das Hangende und Liegende petrographisch resp. durch Fossilien einwandsfrei bestimmt war, derart, daß die fossilleeren Schiefer das obere Mitteldevon, die Buchenauer Kalke und der Pönsandstein das Oberdevon vertreten sollten, und zwar sollte sein der Buchenauer Kalk „unteres Oberdevon“, der Pönsandstein „oberes Oberdevon“. Das oberdevonische Alter der Buchenauer Kalke hielt er deshalb für erwiesen, weil diesen Kalken eine andere etwa 5—6 m mächtige Gesteinslage, die als Tuffbreccie gedeutet wurde, eingelagert sei, welche wiederum neben mitteldevonischen auch oberdevonische Fossilien enthält. Nach einem wiederholten und genauen Studium der Amönauer Steinbrüche muß ich diese Deutung und Beweisführung in vollem Umfange ablehnen. Die Tuffbreccie ist ein echtes Konglomerat, das außerdem den Buchenauer Kalken nicht eingelagert ist, sondern diese überlagert. Außerdem sind

die Buchenauer Kalke mitteldevonisch und reichen höchstens mit dem hangenden Teile in das älteste Oberdevon hinein.

Der Buchenauer Kalk im Sinne Liebers ist recht heterogen. Den untersten Teil der Buchenauer Kalke bilden Schalsteine, wie ich sie in weiter Verbreitung aus der Lahnmulde kenne. Darüber folgt eine Wechsellagerung von dünnen Schalsteinlagen mit dunklen Tentaculiten führenden plattigen Kalken und endlich plattige dunkle Kalke, die mit sandigen Schiefern wechsellagern. In den Steinbrüchen nördlich der Amönauer Landstraße führen die Buchenauer Kalke Kalklinsen, welche Terebratula pumilio enthalten[1]).

Die Plattenkalke entsprechen genau den Plattenkalken, die ich im Sauerlande und auch in der Dillmulde als unzweifelhafte Vertreter der Riffkalke kennen gelernt habe. Dazu kommt, daß die Kalke mit Schalsteinen des Mitteldevon auf das engste verknüpft sind. Sie sind also vom Alter des oberen Mitteldevon und können höchstens bis in das unterste Oberdevon hinaufreichen.

Von besonderem Interesse ist die von H. Lieber als Tuffbreccie gedeutete 5—6 m mächtige Gesteinslage, die in dem kleinen Steinbruche unmittelbar am W-Ausgange von Amönau sehr schön aufgeschlossen ist. Diese Gesteinsbank besteht vorwiegend aus wallnuß- bis faustgroßen Gesteinsfragmenten, zu denen in geringerer Zahl noch größere Gesteinsfragmente hinzutreten. Da der Querschnitt und der Anschliff allein zur Beurteilung dieser Gesteinsfragmente nicht genügte, habe ich mich bemüht, Teile der Gesteinsbank aufzusuchen, aus denen man die Gesteinsfragmente in unverletztem Zustande isolieren kann. Es gelingt das besonders leicht in dem oberen Teile der Gesteinsbank, wo die Gesteinsfragmente weniger fest miteinander verkittet sind. Die so gewonnenen Gesteinsfragmente zeigen vollkommene Abrollung wie Flußkiesel. An der konglomeratischen Natur dieser Gesteinsfragmente kann ein Zweifel nicht mehr bestehen. Es handelt sich also nicht um eine Breccie, sondern um ein typisches Konglomerat. Daß neben vollkommen gerundeten Geröllen sich auch eckige weniger stark abgerollte Gerölle finden, ist selbstverständlich. Sehr instruktiv ist auch ein von H. Lieber angeschliffener Gesteinsblock, der im Querschnitt vollkommen gerundete Komponenten zeigt. Die Angaben in H. Liebers posthumen Werke stimmen mit dem Belegmaterial Liebers nicht vollkommen überein.

Das Material, aus dem das Konglomerat besteht, ist sehr ver-

---

1) Die Untersuchung der Linsen auf ihren Fossilinhalt hin sollen im kommenden Frühjahr fortgesetzt werden.

schiedener Art. Es finden sich als Komponenten typische Riffkalke mit Korallen, Iberger Kalk, Buchenauer Kalke (meist in größeren schwächer abgerundeten Stücken) und endlich stark zersetzte Geröllstücke oberdevonischer Cephalopodenkalke. Es ist nun noch zu beachten, daß unter den Komponenten Gesteine vorhanden sind, die in der weiteren Umgebung von Amönau nicht vorhanden sind und auch verschiedenen Faciesgebieten des oberen Mitteldevon angehören. Eine vulkanische Durchschlagsröhre hätte aber nur Gesteine eines einzigen Faciesgebietes antreffen können. Da außerdem das Muttergestein der Riff- und Cephalopodenkalke im Rimberggebiet nicht vorhanden ist, müßten die Gesteinsfragmente etwa 15—20 km weit transportiert sein. Das ist aber bei einer submarinen Entstehung der Tuffbreccie und bei der bedeutenden Größe einzelner Komponenten wie der Menge des Materials einfach ausgeschlossen.

Im Hangenden der Konglomerate stellen sich Sandsteine und konglomeratische Grauwacken ein, die bereits die Natur des Pönsandsteines besitzen.

Nur bei Nichtbeachtung der Lagerungsverhältnisse kann man weiterhin zu der verfehlten Auffassung kommen, daß die sogenannte Tuffbreccie den Buchenauer Kalken eingelagert ist. In dem Steinbruch an der Amönauer Landstraße sind gefältelte jüngere Buchenauer Kalke auf die nichtgefältelten Sandsteine und darunter liegenden Konglomerate an einer flachen Überschiebung überschoben. Im Hangenden der überschobenen Buchenauer Kalke folgen dann anscheinend nochmals die Konglomeratbänke. Schürfe werden hier die Sachlage klären müssen. Weitere flache Überschiebungen und schichtparallele Verschiebungen sind zurzeit in den Steinbrüchen nördlich der Amönauer Landstraße aufgeschlossen. Insgesamt zeigen meine bisherigen Untersuchungen, daß das Amönauer Devonvorkommen ein Gebiet mit ausgezeichneter Schuppenstruktur darstellt, in dessen nordöstlicher Fortsetzung das gleichgebaute Ensegebiet bei Wildungeu liegt.

Das Alter der Konglomerate kann nicht zweifelhaft sein. Sie bilden die Basis des Fossley und Pönsandsteins. Eine Neuuntersuchung des Langenaubacher Tuffbreccie, die übrigens Tuff nicht zu enthalten scheint, wird in Zukunft nicht zu vermeiden sein.

c) *Das Oberdevon in der Attendorn-Elsper Doppelmulde.*

Diesem interessanten Devonvorkommen hat W. Henke seit dem Jahre 1907 seine besondere Aufmerksamkeit gewidmet. Aus eigener Anschauung kenne ich nur den westlich der Lenne ge-

legenen Teil der Doppelmulde. Zwischen Grevenbrück und Niederhelden zieht als Hangendes der Stringocephalenkalke ein Zug oberdevonischer Kalke entlang. Bei dem Wirtshaus St. Klaas finden sich Nehdener Schiefer mit der Fauna von Nehden. Als häufigsten Goniatiten fand ich auf gemeinsamer Exkursion mit meinem Assistenten Herrn Schindewolf Cheiloceras subpartitum, den W. Henke auffallenderweise nicht erwähnt. Über den Nehdener Schiefern folgen die Cephalopodenkalke, die im Verlaufe des Zuges mehrfach aufgeschlossen sind. Die fossilführenden Kramenzelkalke, aus denen W. Henke östlich des Gehöftes zur Förde eine kleine Fauna gesammelt hat, gehören der Laevigatastufe an. Die hangenden Teile der Cephalopodenkalke, „graublaue Kalkknollen in grauen Tonschiefern" führen nach Henke bei Niederhelden Clymenien „aus der Verwandtschaft der undulata und striata." Henke deutet diese obersten Lagen des Cephalopodenkalkes wohl nur wegen ihrer petrographischen Ähnlichkeit als Wocklumer Kalk. Nach den angeführten Fossilien kann es sich aber nur um Laevigatakalke und zwar um den mittleren Teil handeln. Vorläufig liegt auf Grund der mitgeteilten Fauna auch nicht die geringste Veranlassung vor, für diese Schichten ein anderes Alter in Anspruch zu nehmen. Demnach reichen die Cephalopodenkalke nur bis zur Laevigatastufe hinauf. Rote Tonschiefer sind diesen Cephalopodenkalken nicht eingelagert.

Abweichend ist die Ausbildung des Oberdevon am Südrande der Doppelmulde. Über roten Tonschiefern folgen Schiefer, die mit Sandsteinen wechsellagern. Sehr instruktiv ist das Profil an der Bonzeler Straße. Die Verbreitung der roten Tonschiefer im Gebiete der Attendorn-Elsper Doppelmulde ist nicht ohne Bedeutung. Die Kartierung von W. Henke hat ergeben, daß die roten Schiefer verschieden alten Schichten aufliegen, dem Lenneschiefer, dem Odershäuser Kalk, den Büdesheimer Schiefern und dem Adorfer Kalk. Die Kartierung hat also mit voller Deutlichkeit den Nachweis erbracht, daß die Rotschiefer mit den hangenden Sandsteinen transgredieren. Die Möglichkeit, daß diese mächtige Schichtenfolge, Henkes sandiges Oberdevon, auf kurze Entfernung hin restlos verschwindet und vollständig durch Cephalopodenkalke ersetzt wird, ist m. E. eine Unmöglichkeit. Da weiterhin auch im Gebiete des Cephalopodenkalkzuges stellenweise unter dem transgredierenden Unterkarbon noch Reste des sandigen Oberdevon, insbesondere rote Schiefer, erhalten sind, liegt auch im Attendorn Elsper Devongebiet nur die Möglichkeit vor, daß die Rotschiefer jünger als die Laevigatakalke sind. Demnach sind die roten

Schiefer Fossley und die hangenden Sandsteine Pönsandsteine. Die Unvollständigkeit der Schichtenfolge im Liegenden des Oberdevon führe ich auf Emersionen[1]) zurück.

W. Henke ist zu einer anderen Auffassung gekommen. Nachdem er die Knollenkalke im Hangenden der Cephalopodenkalke nördlich des Himmelsberges als Wocklumerkalk also als oberstes Oberdevon in Anspruch genommen hatte, Cephalopodenkalke und Knollenkalke somit das Oberdevon nach seiner Ansicht ganz ausfüllten, mußte er notwendiger Weise zu der Anschauung gelangen, daß das sandige Oberdevon den Cephalopodenkalken altersgleich zu setzen sei, mithin diese vertritt.

Diese Anschauung läßt sich nicht halten, da die Cephalopodenkalke nur bis zur Laevigatastufe hinaufreichen. Außerdem ist eine derartige Vertretung so mächtiger Schiefer und Sandsteine in einem so eng umgrenzten Gebiete, wie es der südwestliche Teil der Elsper Mulde darstellt — Entfernung zwischen sandigem und kalkigem Oberdevon 1 km — m. E. einfach eine Unmöglichkeit.

Die Fossleytransgression verbunden mit der heute nicht mehr zu bestreitenden Transgression des obersten Unterkarbons (Culm) führen zu einer einfachen und klaren Sachlage, die nicht mit merkwürdigen und unerklärlichen Faciesveränderungen zu rechnen braucht.

3. Begriff und Deutung des Pönsandsteins und des Fossley.

Die sandige Ausbildung des oberen Oberdevon umgürtet den Siegerländer Block. Sie ist in sicherer Lagerung in der Dillmulde, im Rimberggebiet, im Dreiherrensteingebiet, im Kellerwalde und in der Attendorn-Elsper Doppelmulde nachgewiesen. Den Nordrand des Rheinischen Gebirges erreicht die sandige Ausbildung nicht. Im Briloner Devongebiet sind Tonschiefer vorhanden, die gegen W. hin kalkreich werden. Das sandige Material kann also nicht von N und auch nicht von O stammen. Die heutige Verbreitung weist unmittelbar darauf hin, daß es von der Siegerländer Insel resp. Halbinsel herstammt. Vermutlich handelt es sich um abgetragene Koblenzschichten. Demnach bezeichne ich als Pönsandstein diejenigen sandig-schiefrigen Schichten des obersten Oberdevon (Stufe VI), deren Material, aus der Abtragung des Siegerländer Blockes

1) Auf Emersion führe ich auch die konglomeratische Ausbildung der obersten Massenkalkbänke am Rehnert zurück.

herstammend, in der weiten Umrandung des Siegerländer Blockes sedimentiert wurde. Aus der Mächtigkeit und Verbreitung der Pönsandsteine geht nunmehr auch ohne weiteres hervor, daß sie nicht gleichartig mit den Cephalopodenkalken sein können. Es ist m. E. unmöglich, daß die Pönsandsteine in der Richtung der Sedimentationszufuhr plötzlich abbrechen, durch Cephalopodenkalke vertreten werden, dann ebenso unvermittelt wieder hervortreten, und neuerdings wieder durch Cephalopodenkalke ersetzt werden können.

Eine vom Pönsandstein scharf zu unterscheidende Gesteinsfolge ist der Fossley, die Unterlage des Pönsandsteins. Seiner Natur nach ist der Fossley ein vorwiegend rot gefärbter Tonschiefer, der den örtlichen Sedimentationsverhältnissen entsprechend mit Grünschiefern wechsellagern kann. Ich habe ihn als horizontbeständige Gesteingruppe aufgefaßt und befinde mich in dieser Hinsicht in Übereinstimmung mit dem Brauch der praktischen Geologen, die Sedimente gleichen petrographischen Charakters so lange als gleichaltrig betrachten, so lange nicht palaeontologische Gründe zu einer abweichenden Behandlung derartiger Sedimente zwingen. In der Fossleyfrage weichen sie freilich von ihrem Brauche ab.

Ahlburg[1]) hat jüngst die Anschauung ausgesprochen, daß die Rotfärbung der Rotschiefer als vulkanische Nachwirkung submariner Diabaseruptionen aufzufassen sei. Dem gegenüber ist zu betonen, daß der Fossley auch in Gebieten auftritt, in denen seit der mitteldevonischen Zeit keine Eruptionen mehr stattfanden und daß die oberdevonischen Diabase jünger sind als der Fossley. Es ist von besonderer Bedeutung, daß die Rotschiefer des oberen Unterdevon, Mitteldevon und Oberdevon streng an den nordatlantischen Kontinent gebunden sind, und gleichsam eine stärkere oder auch sandige Sedimentation einleiten. Mit der Entfernung vom nordatlantischen Kontinent keilen die Rotschiefer aus.

Den mir bisher bekannten Tatsachen würde folgende Deutung des Fossley entsprechen:

1. Während der Ruheperioden in der Abtragung des nordatlantischen Kontinentes findet eine intensive Verwitterung der auf dem nordatlantischen Kontinente anstehenden Gesteine statt. Die Verwitterungsprodukte sind rot gefärbt. In dem abhängigen Sedimentationsgebiete bilden sich Kalke oder die diese vertretenden dunklen Tonschiefer.

---

1) Zeitschrift für praktische Geologie, 25. Band, 1917, S. 29 ff.

2. Bei Beginn einer neuen stärkeren Abtragung des nordatlantischen Kontinentes werden zunächst die rotgefärbten Verwitterungsprodukte abgespült. Dann greift die Erosion tiefer. In dem abhängigen Sedimentationsgebiet bilden sich demnach zunächst rotgefärbte Tone, dann Sandsteine oder Grauwacken und Tonschiefer. Die roten Schiefer müssen daher mit der Entfernung vom nordatlantischen Kontinente verschwinden.

Demnach bezeichne ich als Fossley bunte vorherrschend rote, vorwiegend vom nordatlantischen Kontinente[1]) herstammende, umgelagerte Verwitterungsprodukte, die die Basis des Pönsandsteins bilden und jünger sind als die Laevigatastufe. Da der Fossley fossilarm ist und nur planktonische Fossilien führt, ist es einstweilen nicht möglich, die ihm entsprechende Zeitspanne auch biostratigraphisch zu definieren.

Fossleyähnliche Bildungen sind vor allem oberdevonische Rotschiefer, die älter sind als Fossley und daher nicht mit dem gleichen scharf definierten Namen belegt werden sollten. Am Nordrande des rheinischen Gebirges erfahren die Cephalopodenkalke von O gegen W hin eine vollständige facielle Umprägung, sodaß im W die Schichtenfolge des Oberdevon auf neuer Grundlage zu erarbeiten ist. W. Paeckelmann hat das in seiner Studie über „das Oberdevon des Bergischen Landes“ versucht. Das Oberdevon ist in diesem Gebiete vorwiegend als Rotschiefer ausgebildet, die er als Cypridinenschiefer bezeichnet. Außerdem wird die oberdevonische Schichtenfolge von ihm als vollständig oder lückenlos bezeichnet. Die unteren Cypridinenschiefer gehören der Stufe II an, die mittleren Cypridinenschiefer enthalten keine Goniatiten. Die oberen Cypridinenschiefer hat W. Paeckelmann als Schichten vom Alter der Wocklumer Kalke bezeichnet. Die von mir ausgeführte Neubestimmung dieser Clymenien ergab das Resultat, daß es sich um Clymenien der Gruppe der Clymenia laevigata handelt. Bisulcate Clymenien fehlen in dem gesamten Material vollkommen. Auch diese würden kein Wocklumer Alter beweisen. Durch Beobachtung vereinzelter Lobenlinien konnte diese Bestimmung einwandfrei durchgeführt werden. Auch die übrigen Clymenien entsprechen

---

1) Es liegt naturgemäß nahe, das Material der Rotschiefer auch von der mitteleuropäischen Insel herzuleiten. Einstweilen widersprechen dem eine Reihe von Beobachtungen, so z. B. die Beobachtung Ahlburgs, daß der Fossley an den Rändern des Beckens in dunkle Tonschiefer übergeht. — Bezüglich der Deutung der Rotschiefer vergleiche man auch R. Richter: Zur Stratigraphie und Tektonik der Ösling-Eifel-Mulde.

nach dem Verlauf der Anwachsstreifen ganz den Clymenien der Tiefzone der Laevigatastufe. Da das Oberdevon des Bergischen Landes somit von unten her nur bis zur Laevigatastufe hinaufreicht, und dann sofort Unterkarbon folgt, kann von einem lückenlosen Profile nicht die Rede sein. Die petrographische Kontinuität der Sedimente täuscht hier ganz entschieden.

Die Rotschiefer können, da sie eine andere stratigraphische Position einnehmen, nicht als Fossley bezeichnet werden. Die Einflüsse des nordatlantischen Kontinentes, die sich bereits in der Gegend von Iserlohn durch das Auftreten von Rotschiefer unterhalb des Fossley bemerkbar machen, treten im unteren und mittleren Oberdevon des Bergischen Landes viel stärker hervor. Die Abspülung der roten Verwitterungsprodukte war während des unteren und mittleren Oberdevon von geringerer Intensität, und daher ganz auf die küstennahen Meereszonen am nordatlantischen Kontinente beschränkt. Ich nenne sie im Gegensatz zum Fossley „die roten Randschiefer".

4. Das Gesamtbild.

Den Sedimenten des altrheinischen Sedimentationscyclus, hervorgerufen durch die Abtragung des nordatlantischen Kontinentes und der mitteleuropäischen Insel, sind im oberen Mitteldevon Flinzkalke mit den randlichen Riffkalken aufgelagert. Die Flinzkalke, hier im weiteren Sinne gebraucht, bilden die zentrale Ausbildung. Das ist die Unterlage des Oberdevon. Die oberdevonischen Cephalopodenkalke bilden die jüngsten und Scheitelschichten des altrheinischen Sedimentationscyclus. Die Grenze der Riff- und Flinzkalkplatte stellt entweder eine Verzahnung der beiden mitteldevonischen Facies, Flinz- plus Riffkalk, mit den oberdevonischen Cephalopodenkalken dar, oder ist infolge einer Emersionsdiskordanz scharf und unvermittelt. Die oberdevonischen Cephalopodenkalke stellen eine den Orthoceraskalken Skandinaviens vergleichbare Kalkplatte dar, die von Thüringen und dem Harz aus sich in die Buchten des rheinischen Beckens hinein erstreckt, und mit der Annäherung an die Küsten des Siegerländer Blockes und des nordatlantischen Kontinentes eine Faciesveränderung erleidet. In der Regel gehen die Kalke in eine tonige, an der Küste des nordatlantischen Kontinentes in die Rotschieferfacies über.

In dem unteren Teil der Cephalopodenkalkplatte greift dreimal eine Schieferfacies ein, die die Cephalopodenkalke verdrängt. Die Zone II α ist bei Nehden, Nuttlar und Attendorn durch die

vollkommen gleichartigen Cephalopodenschiefer ersetzt. In der gleichen Weise ist in weiter Verbreitung die Zone I δ besonders linksrheinisch durch Cephalopodenschiefer vertreten. Das sind die Büdesheimer Schiefer oder Schiefer von Matagne. Es ist sehr auffallend, daß in der gleichen stratigraphischen Position die Bickener und Kellwasser Kalke mit ihrem Schieferreichtum hervortreten. Endlich kann auch ein tieferer Teil der Stufe I durch Cephalopodenschiefer vertreten sein. Die Verbreitungsgebiete dieser Cephalopodenschiefer sind noch nicht hinreichend bekannt, um ihre Natur zu verstehen. Nur bei den liegenden Schiefern handelt es sich wohl um ein Hinaufreichen einer modifizierten Flinzfacies ins Oberdevon.

Die Zeitspannen, während derer sich die Cephalopodenkalke bildeten, bedeuteten für die Kontinente Zeiten geringster Abtragung aber starker Verwitterung der anstehenden Gesteine. In den ruhigen Ablauf dieser Vorgänge greifen noch vor Beginn der karbonischen Zeit störend epeirogenetische und orogenetische Vorgänge ein. Die orogenetischen Bewegungen bedeuten Faltung des Meeresbodens. Längs der Sattellinien findet eine Abtragung der oberdevonischen, zum Teil auch der mitteldevonischen Sedimente statt. Im Zwischengebirge [1]) östlich des Siegerländer Blockes werden die Cephalopodenkalke restlos abgetragen. Hier wird nur sandiges Oberdevon angetroffen.

Nach der Faltung macht sich eine langsame und stärkere Hebung des Siegerländer Blockes und des nordatlantischen Kontinentes bemerkbar. Die Folge ist der neurheinische Sedimentationscyclus. Der Fossley leitet diese Neuphase ein. Der Fossley transgrediert über die abgetragenen Sättel. Dann ruft die Abtragung des Siegerländer Blockes die Bildung des Pönsandsteins hervor. Die Scheitelschichten des neurheinischen Sedimentationscyclus stellen die Cephalopodenkalke nebst den Brachiopoden- und Crinoidenkalken des älteren]Unterkarbon und die Wocklumer Kalke des nördlichen Sauerlandes dar. Die kalkige Facies erscheint demnach im Innern des Gebirges später als am Außenrande. Dem Altunterkarbon gehören die Kalke von Erdbach-Breitscheid an, deren Fauna auf das engste mit der Wocklumer Fauna verknüpft ist. Auf die Wocklumeriastufe des Oberdevon folgt die Pericyclusstufe und erst dann die eigentliche Glyphiocerasstufe (oder

---

1) Als Zwischengebirge bezeichne ich die östlich des Siegerländer Blockes zwischen dem sauerländischen und hercynischen (Dill- und Lahnmulde, Kellerwald und Nordharz) Faltungsgebiet gelegene tektonische Einheit.

Culm). Culm ist, wie meine fortgesetzten Studien über das Unterkarbon gezeigt haben, junges Unterkarbon[1]). Der Culm transgrediert mit den Alaunschiefern oder auch mit sandigen Bildungen oder Konglomeraten. Die Lücke im Liegenden des Culm ist eine größere, als ich bisher angenommen hatte. Die beiden Transgressionen, die Fossley- und die Culmtransgression bedingen die Merkwürdigkeiten der oberdevonischen Schichtenschnitte.

1) Ausführlich wird Herr Brüning, der die Untersuchungen meines verstorbenen Freundes K. Burhenne fortsetzt, das Unterkarbon in einer Dissertation behandeln.

Über die Ausgestaltung der Lobenlinie bei Jura- und Kreide-Ammonoideen.

Ein Beitrag zur Frage der Entwicklungsrichtung und Organisationshöhe.

Von

Hans Salfeld, Göttingen.

Vorgelegt von Herrn Stille in der Sitzung vom 31. Oktober 1919.

1913 habe ich schon darauf hingewiesen, daß es nach unseren bisherigen Erfahrungen im Jura und der Kreide nur vier Ammonoideenstämme gibt: die *Phylloceratida* [1]), *Lytoceratida*, *Stephanoceratida* und *Harpoceratida* (Salfeld, Die zoogeographische Stellung des süddeutschen Jura. Zeitschr. d. deutsch. geol. Ges.). Auf Grund der Richtung der Zuwachsstreifen oder den diesen gleichsinnig gerichteten Rippen lassen sich diese Stämme nicht unterscheiden, wenn auch die *Harpoceratida* vorwiegend sichelförmige Zuwachsstreifen oder Rippen besitzen, so ist dies doch nicht durchgehends der Fall, wie auch bei den *Phylloceratida*, *Lytoceratida* und *Stephanoceratida* neben vorwiegend geraden Zuwachsstreifen oder Rippen typisch sichelförmige vorkommen. Es sind gewisse Charaktere der Lobenlinie und der sekundären Skulpturen (Berippung, Beknotung usw.), welche gestatten, die Stämme festzulegen. Doch ist die Vielseitigkeit dieser Charaktere nicht in kurze Formeln zusammenzufassen, da alles der Entwicklung und auch der Hemmung unterliegt, und echter Atavismus das Bild verschieben kann.

Ich beschränke mich hier auf die Festlegung der Charaktere des Stammes der *Phylloceratida* nach dem mir vorliegenden Materiale. Der Stamm der *Stephanoceratida* ist auf die gleiche triadische Wurzel wie die *Phylloceratida* zurückzuführen, auf die *Monophyllen*, schlägt aber in der Ausbildung der Lobenlinien eine andere Entwicklungsrichtung ein. Aus den *Stephanoceratida* entwickeln sich

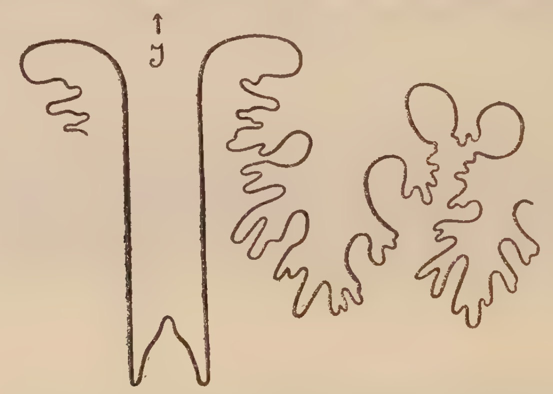
1) Die vier Stämme würden den Rang von Superfamilien einnehmen, da ich dem Brauche der Zoologen folgend, die Cephalopoden als Ordnung, die Ammonoidea als Sektio betrachte.

sicher die *Harpoceratida*, wahrscheinlich auch die *Lytoceratida* im unteren Lias. Diese Fragen können nur gestreift werden. Im übrigen behandele ich die Ausbildung gewisser Charaktere der Lobenlinie in ihrem zeitlichen Auftreten in den verschiedenen Stämmen.

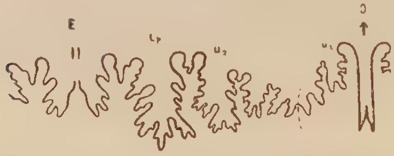
Phylloceratida.

Nach der Ausbildung der Sattelendigungen der Lobenlinien werden oft die *Phylloceratida* als Formen mit macrophyll zerschlitzten Sattelendigungen den drei übrigen Stämmen als Formen mit microphyll zerschlitzten Sattelendigungen gegenübergestellt. Demgegenüber läßt sich feststellen, daß ein großer Teil der *Phylloceratida* der Kreide, besonders die Gruppe *Schlüteria* des Senon, typisch microphyll zerschlitzte Sattelendigungen aufweisen. Wir müssen daher nach anderen Merkmalen suchen, wenn auch zuzugeben ist, daß die jurassischen *Phylloceraten* macrophyll zerschlitzte Sattelendigungen besitzen.

Die Ausbildung des Innen- oder Antisiphonallobus (J) bei *Phylloceras* und *Rhacophyllites* weist, neben der primären Zweispitzigkeit, keine weitere Zerschlitzung auf (Textfig. 1 u. 2). Dieser



Figur 1. Phylloceras Zignodianum d'Orb. Oxfordien.



Figur 2. Rhacophyllites eximius Haner.

alte Charakter, der ihnen von den triadischen Vorläufern, den *Monophylliten* (einschl. *Megaphyllites*), überkommen ist, wird bis in die obere Kreide hinein zäh festgehalten (*Schlüteria velledaeformis*

Schl. aus dem Senon von Lüneburg. Taf. I, Fig. 6). Daneben sehe ich die Form des gerundeten, am Grunde stark eingeschnürten und gestielten Blattes, welche die Sattelendigungen aufweisen, als charakteristisch für die *Phylloceratida* an, wie dies auch von den meisten Autoren angegeben wird. Als zweiter alter Charakter tritt uns in der Ausgestaltung der Lobenlinie die mehr oder weniger hohe, aber stets dachförmige Gestalt des Mediansattels im Außenlobus (E) entgegen. Hierauf pflegte mein Lehrer Pompeckj aufmerksam zu machen.

Typ der reinen Reihe.

Für die reine Reihe der *Phylloceratida* ist demnach folgende Diagnose aufzustellen:

Zuwachsstreifen linear bis mehr oder weniger deutlich sichelförmig. Lobenlinie nach der Formel E Lp U_2 U_3 (= S) U_1 J gebaut (siehe Salfeld, Monographie der Gattung *Cardioceras*. Zeitschr. d. deutsch. geol. Ges. 1915, S. 151). J zweispitzig, lituid[1]), ohne weitere Zerschlitzung der Seitenwände, J mit dachförmigem Mediansattel, Sattelendigungen macrophyll (Jura bis Unterkreide) bis microphyll (Unterkreide bis Oberkreide), gerundete, am Grunde eingeschnürte und gestielte Blätter.

Die sämtlichen Sättel der Innenlobenlinie sind monophyll oder,

---

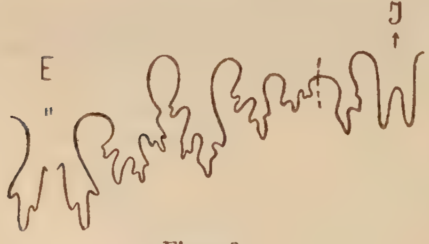
1) Die Form des Innenlobus ist für die reine Reihe der *Phylloceratida* und einiger bisher festgestellter Nebenreihen außerordentlich charakteristisch. Dies ist schon von Quenstedt, der den Innenlobus goniatitisch nennt, und von Neumayr erkannt, ohne daß von beiden auf das Wesentliche der Form und ihrer Bedeutung für die stammesgeschichtlichen Zusammenhänge eingegangen ist. Der Innenlobus zusammen mit dem anschließenden Sattel erinnert an das Zeichen des Widders, noch treffender wäre der Vergleich mit zwei spiegelbildlich zu einander gestellten Krummstäben, wenn wir die mehr oder weniger reich zerschlitzten Seiten von U_1 und die immer gegen außen gerichtete Ausladung des ersten, immer ganzrandigen Innensattels hinzunehmen. Um für dies Bild einen kurzen Ausdruck zu geben, bezeichne ich diesen Typ der Innenlobenlinie (bezogen auf Innenlobus und anschließenden Sattel) als lituid (lituus = Krummstab).

Unter allen triadischen Ammonoideen finden wir außer der überall vorhandenen Zweispitzigkeit das Fehlen jeder weiteren Zerschlitzung des Innenlobus (abgesehen von manchmal in Erscheinung tretenden ein oder zwei kleineren Seitenzacken) nur noch bei den *Ceratiten* im weitesten Sinne. Damit will ich hier keine näheren Zusammenhänge zwischen den *Ceratiten* und den *Phylloceratida* konstruieren. Bei den übrigen triadischen Ammonoideen ist der Innenlobus und der anschließende Sattel mehr oder weniger reich zerschlitzt und gewährt in dieser Beziehung dasselbe Bild, wie es die *Ammonoideen* der Jura- und Kreidezeit unter Ausschluß der *Phylloceratida* bieten. Siehe z. B. den Innenlobus von *Gymnites* aus dem alp. Muschelkalk auf Taf. I, Fig. 1.

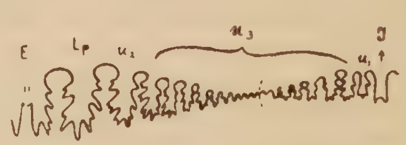
wie NEUMAYR (Jurastudien, Jahrb. k. k. geol. Reichsanst, XXI. Bd., 1871) gezeigt hat, ist der zweite Innensattel bei allen jenen Formen zweiblätterig, welche auf dem Steinkerne radiale ununterbrochen über Flanken und Außenseite laufende Einschnürungen tragen (Fextfig. 1). Als Ausnahme hiervon kannte NEUMAYR *Phyll. tortisulcatum* d'ORB., dem ich *Phyll. Guettardi* d'ORB. und *numismale* (QU) POMP. hinzufüge. Ferner konnte ich für *Phyll. silesiacum* OPP. aus dem Obertithon feststellen, daß auch der dritte Innensattel zweiblättrig ist. Bei der Formenreihe des *Phyll. tatricum* PUSCH habe ich weiter die normalen monophyllen Innensättel festgestellt.

Die Wurzel der *Phylloceratida*.

ZITTEL trennt im Handbuch der Palaeontologie Bd. II die enggenabelten Formen als *Phylloceras s. str.* von den weitgenabelten ab, welche er als *Rhacophyllites* bezeichnet. Als Stammformen von *Phylloceras* sieht er die ebenfalls enggenabelten triadischen *Megaphyllites* an, monophylle Formen, deren Sättel sämtlich in einem am Grunde eingeschnürten Blatt endigen, während er *Rhacophyllites* an die ebenfalls weitgenabelten *Monophyllites* der mittleren und oberen Trias anschließt. Da wir wissen, daß weit- und enggenabelte Formen von *Ammonoideen* oft als Plus- und Minusabweicher zu ein und derselben Art gerechnet werden müssen, bleibt dies Kriterium allein immer von sehr problematischer Natur. Besser ist schon die von SMITH in der zweiten Auflage von ZITTEL-EASTMAN vorgenommene Trennung auf Grund der Lobenlinien; *Megaphyllites* mit cycloloben, d. h. kurzen, mehr oder weniger kreisförmigen Sattelendigungen, und *Monophyllites* mit mehr oder weniger löffelförmigen Sattelendigungen. Dies trifft aber m. E. noch nicht das Wesen der Differenz; die Sattelwände bei *Monophyllites* sind asymetrisch, nur einseitig zerschlitzt (Textfig. 3), bei *Megaphyllites* symetrisch, zweiseitig zerschlitzt (Textfig. 4). *Monophyllites* ist der



Figur 3.
Monophyllites sphaerophyllus Hauer.



Figur 4.
Megaphyllites Jarbas Münster.

ältere Typ, aus dem sich *Megaphyllites* entwickelt hat, da sich dieser Weg verfolgen läßt.

Bei *Phylloceras* und *Rhacophyllites* zeigt der Innenlobus und anschließende Sattel die ursprünglichste Form von *Monophyllites*, die sonstigen Innensättel entweder eine etwas vorgeschrittene monophylle Form oder ein *Megaphyllites* Typ. Hieraus und aus dem im weiteren angeführten Beweisen läßt sich der Satz aufstellen:

Bei den Ammonoideen der Jura- und Kreidezeit ist die Innenlobenlinie, besonders der Innenlobus der Sitz primitiver Charaktere und gestattet stammesgeschichtliche Zusammenhänge zu entschleiern.

Über die Trennung von *Rhacophyllites* herrscht völlige Unklarheit. Zittel sieht die Weitnabligkeit als brauchbaren Charakter an gegenüber den enggenabelten *Phylloceras*-Formen. Nach der Ausbildung der Lobenlinien lassen sich die meist weitgenabelten Formen der oberen Trias als diphylle *Rhacophylliten* von den tri- oder tetraphyllen *Phylloceraten* des Jura und der Kreide trennen. In der ontogenetischen Entwicklung der Lobenlinien dieser beiden Gruppen treten die Unterschiede deutlicher hervor. *Rhacophyllites* besitzt in der Jugend eine monophylle Lobenlinie vom *Megaphyllites*-Typ, in späteren Wachstumsstadien bildet sich im Sattel eine mediane Incision, welche schnell die beiden primären Seitenincisionen überflügelt (Tafel I, Fig. 18). Diese beiden Seitenincisionen kann ich nur der α- und β-Incision von Dietz (Dietz, Über bipolare Lobenzerschlitzung. Centralbl. f. Min. etc., 1916, S. 195) gleichsetzen und damit auch den beiden Seitenincisionen unterhalb des Sattelblattes bei *Megaphyllites:* bei *Monophyllites* würde normal nur die α- oder β-Incision vorhanden sein. Diese Unterschiede sind gradueller Natur und zahlreiche Übergangsformen zeigen, daß der Entwicklungsgang von *Monophyllites* über *Megaphyllites* zu *Rhacophyllites* führte. Deshalb vermag ich auch nicht den Schnitt zwischen *Mesoammonoideen* und *Neoammonoideen* im Sinne Wedekinds zu führen, der die *Monophyllen* als monopolar ansieht, denn auch bei diesen ist in der ontogenetischen Entwicklung eine bipolare Zerschlitzung festzustellen, wenn auch die α- und β-Incision an den Seiten des Sattels auftreten, also etwas tiefer als bei den jüngeren *Phylloceraten* aus Jura und Kreide, wo die α- und β-Incision auf den Sattel hinauf gerückt ist. Die letztere Ausbildung findet sich erstmalig bei *Amm. planorboides* Gümb. aus dem alpinen Rhät. und wird bei allen Jura- und Kreide-Ammonoideen im Grundplan festgehalten. Man kann daher die Jura-Kreide-Ammonoideen als bipolare „*Pachycampyli*" zusammenfassen, weil die jugendliche

Anlage der Sättel breit (Tafel I, Fig. 4) ist, im Gegensatz zu den bipolaren triadischen *Phylloceratida*, deren jugendliche Sattelanlagen (Tafel I, Fig. 18) schmal sind („*Leptocampyli*“). Diese Bezeichnungen decken sich nicht mit den von Hyatt gebrauchten, der sie auf das fertige Lobenbild bei Jura-Kreide-Ammonoideen anwendet (Zittel-Eastman, Textbook of Palaeontologie, I. Aufl.)

Bei *Rhacophyllites* wird gegenüber *Megaphyllites* ein neuer Charakter, die Scheitel- oder γ-Incision, adordiniert. Hierdurch kommen die diphyllen Sattelendigungen zustande. Bei *Phylloceras* werden die primären Sattelendigungen triphyll, dadurch, daß die α- und β-Incision auf den Sattel hinaufrückt. In manchen Reihen wird aber eine γ-Incision adordiniert, dann sind die Sättel von primär tetraphyllem Bau. Die beiden Gattungen so charakterisiert, ist *Phylloceras* auf Jura und Kreide beschränkt, einerlei ob die Formen weit- oder enggenabelt sind, und *Rhacophyllites* auf die obere Trias, auch gleichgiltig, ob die Formen weit- oder enggenabelt sind.

In *Rhacophyllites* ist auch die Hyattsche „Gattung“ *Discophyllites* einzubeziehen, da die einzige Art, *Amm. patens* Mojs. sich von den typischen *Rhacophylliten* nur durch den monophyllen Außensattel unterscheidet. Wir dürften es hier mit einer Hemmungserscheinung zu tun haben, da nach unseren Erfahrungen zuerst der Außensattel der Zerschlitzung unterliegt. Ganz sicher gehört aber *Phyll. larianense* Mengh. aus den Margaritatus-Schichten des mittleren Lias nicht zu *Rhacophyllites*, sondern nach der Ausbildung der Lobenlinie als Form mit hahnenkammartigem Kiel zu der Gruppe des *Phylloceras tortisulcatum* d'Orb. Es ist m. E. überflüssig für diese eine Art die Hyattsche Gattung *Meneghiceras* aufrecht zu erhalten. Taf. I, Fig. 9.

Ob die Gattung *Euphyllites* Wähner, die Gruppe des *Phyll. Bonifaci* Wähn. aus dem unteren alpinen Lias, sicher zu *Phylloceras* gehört, ist erst nach dem Studium der Innenlobenlinie zu entscheiden. Die Lobenlinie der Außenseite spricht jedenfalls dafür.

Ob sich *Phylloceras* von den älteren *Rhacophylliten* ableitet oder direkt von *Megaphyllites* oder etwa von *Mojsvarites* der Gruppe des *Amm. planorboides* Gümb., ist z. Zt. noch nicht zu entscheiden. Letzteres würde dadurch an Wahrscheinlichkeit gewinnen, weil mit dieser rhätischen Form erstmalig der jurassische primär dreilappige Sattel auftritt, also eine neue Entwicklungsrichtung eingeschlagen wird, welche auch *Phylloceras* eigentümlich ist. Die Ausbildung des Innenlobus von *Mojsvarites* ist uns aber noch unbekannt, damit sind auch alle weiteren Spekulationen müßig.


Tafel I
1
Gymnites Palmai Hauer
2
Ph. numismale Qu.
h. Loscombi
5
d'Orb.
3a
3b
4
6
Tragophyll. ibex Qu.
Schlüt. velledaeformis
7
8
Rh. neojurensis
h. semisulcatum
9
Ph. lariense
11
Paraphyll. calliphyllum
10
Psil. laeve
Pl. placenta
18
12
13
14
Coilopoceras
19
Placent. whitfieldi
Pl. syrtale
15
Engenoceras
16
17
Sphen. lenticularis
Sph. Beecheri

Daß die *Monophyllen* wahrscheinlich auf die *Cycloloben* der älteren Trias und des Permo-Karbon zurückgeführt werden müssen, ist immer behauptet worden, und dürfte auch sehr wahrscheinlich sein. *Cyclolobus* selbst gehört jedoch einem Seitenzweig an, da seine Lobenlinie gegenüber *Monophyllites* höher organisiert ist und demnach nicht die Stammform sein kann.

Abgezweigte Reihen von *Phylloceras*.

Zu verschiedenen Malen kommt es unter den *Phylloceraten* zu einer Weiterentwicklung des Innenlobus:

1. im mittleren Lias bei

Tragophylloceras Hyatt, der Gruppe des *Phylloceras ibex* Qu., *numismale* Qu. und *Loscombi* d'Orb., Formen, welche zu stärkerer Skulptur und einem flach-scheibenförmigen Gehäuse neigen. Von den reinen Reihen sind diese dadurch unterschieden, daß die Seitenwände des Innenlobus eine beginnende Zerschlitzung aufweisen. *Phyll. Loscombi* zeigt eine kleine Seitenzacke (Tafel I, Fig. 5). Bei *Phyll. numismale* ist diese Seitenzacke deutlicher ausgeprägt; es tritt unterhalb derselben eine Auswölbung, eine beginnende zweite Zacke, hinzu (Tafel I, Fig. 2). *Phyll. ibex* weist eine tief liegende Seitenzacke auf und die beiden Spitzen des zweispitzigen Innenlobus drängen schräg gegen außen. Bei einem Durchmesser von ca. 25 mm ist oberhalb der ersten eine kleine zweite Seitenzacke hinzugetreten (Tafel I, Fig. 3):

2. im Turon und Senon, bezw. Cenoman oder bereits Albien:

a) *Engenoceras* Gabb (Tafel I, Fig. 15) aus dem Albien-Cenoman, einschl. *Protengenoceros* Hyatt und *Metengenoceras* Hyatt (Hyatt, *Pseudoceratites* of the Cretaceous. U. S. Geol. Surv. Mongr. XLIV) sind wegen der Ausbildung des Innenlobus noch am ehesten als Hemmungsformen zu den *Phylloceratida* zu stellen. Die extrem hochmündigen Formen vermehren ihre Loben wie bei *Sphenodiscus* dadurch, daß die drei Hauptzacken des Lp zu selbständigen Loben werden, indem die zwischen liegenden Stücke zu selbständigen Sätteln hochwachsen. Weiter werden auch die beiden Primärincisionen des Außensattels zu mehr oder weniger selbständigen Loben.

b) *Sphenodiscus* Meek aus der oberen Kreide (Turon?, Senon) weist den für *Phylloceras* typischen ersten Innensattel auf, die Seitenwände des Innenlobus besitzen eine verschieden weitgehende Zerschlitzung (Tafel I, Fig. 16—17). Während in dieser Beziehung eine Weiterbildung stattgefunden hat gegenüber den Charakteren

der reinen *Phylloceraten*reihen, zeigt die Ausgestaltung der Sättel der Außenseite eine mehr oder weniger weitgehende Hemmung und Umgestaltung, sodaß die Sättel zu je einem großen, am Grunde stark eingeschnürten Blatt werden, das an seinem Oberrande gezackt ist. Eine Lobenvermehrung tritt wie bei *Engenoceras* dadurch ein, daß die drei Hauptzacken des *Lp* zu selbständigen Loben werden.

c) Mit *Sphenodiscus* in engster Beziehung scheint *Placenticeras* (Tafel I, Fig. 12, 13, 19) der oberen Kreide zu stehen. Die Innenlobenlinien, welche Hyatt (l. c.) abbildet, zeigen gegenüber *Sphenodiscus* eine weitergehende Zerschlitzung nicht nur inbezug auf den Innenlobus, sondern auch die Innensättel, welche bei *Sphenodiscus* noch sämtlich einblättrig sind, besitzen bis auf den ersten Innensattel weitergehende Zerschlitzung. Bei *Placenticeras placenta* Deak. kommt es schließlich auch zu einer Zerschlitzung des ersten Innensattels (Tafel I, Fig. 19). Wären nicht Übergänge hierzu in den übrigen *Placenticeras*-Arten vorhanden, so würde man schwerlich dazu kommen, diese Art bei den *Phylloceratida* unterzubringen.

d) In diesem Zusammenhang liegt es nahe, die Gattung *Coilopoceras* Hyatt aus dem Turon auch zu den *Phylloceratida* zu stellen (Tafel I, Fig. 14). Die Sattelblätter der äußeren Lobenlinie sprechen entschieden dafür und nach Analogie der zu *Placenticeras* gestellten Formen die Zerschlitzung des ersten Innensattels nicht dagegen.

e) Während bei allen den vorgenannten Gruppen die Endigung des Innenlobus wie bei *Phylloceras* zweispitzig war, wird von Solger für *Neoptychites* Kossmat aus dem Turon ein einspitziger Innenlobus abgebildet (Solger, Fossilien der Mungokreide in Kamerun, aus E. Esch, Beiträge zur Geologie von Kamerun, 1904, Fig. 18a). Ich habe dies bisher nicht nachprüfen können, doch sehen wir später in anderen Entwicklungsreihen, daß die Tendenz bei den Jura-Kreide-Ammonoideen ganz allgemein vom zweispitzigen zum einspitzigen Innenlobus führt. Die sonstige Ausgestaltung der Lobenlinie stellt jedenfalls *Neoptychites* in allernächste Nähe von *Sphenodiscus* und *Placenticeras*, worauf schon von anderer Seite hingewiesen wurde.

f) Das gleiche trifft für die gleichaltrige Gruppe *Hoplitoïdes* v. Koenen zu, deren Innenlobenlinie bisher nicht bekannt geworden ist, und die ich auch an dem im Göttinger Institut vorhandenen Originalmateriale nicht frei legen kann, ohne die Stücke zu stark zu beschädigen. Übrigens vermag ich keinerlei Konvergenz in den beiden Lobenbildern festzustellen, welche Solger (l. c.) Fig. 74 von *Hoplites* (besser *Leopoldia*) *Leopoldinus* d'Orb. und *Hoplitoïdes Wohlt-*

manni v. K. nebeneinander stellt. Im allgemeinen kann man sagen, was man nicht definieren kann, sieht man als *Hoplites* an, wenigstens wenn es in der Kreide vorkommt.

Es mag immerhin möglich sein, daß wir bei kritischer Verfolgung von Entwicklungsreihen, einen Teil des sog. Kreide-Ceratiten, außer den oben aufgeführten, noch bei den *Phylloceratida* unterzubringen haben.

Schlussfolgerungen.

Bisher ist also festgestellt, daß es bei der Ausgestaltung der Innenlobenlinie einmal im mittleren Lias und (vielleicht mehrfach) in der oberen Kreide bei den *Phylloceratida* zur Zerschlitzung der Seitenwände des Innenlobus, wie auch zur Zerschlitzung des ersten Innensattels in der oberen Kreide kommt. Der bisher einwandfrei festgestellte erste Versuch des Überganges zur Zerschlitzung des Innenlobus bei *Tragophylloceras* des mittleren Lias scheint nicht zu einer längeren Reihe von Formen geführt zu haben, wie dies in der oberen Kreide der Fall gewesen ist. Ich möchte aber auf den ersten Fall nicht den von Wedekind (Beiträge zur progressiven Entwicklung der Organisation der Organismen. Ges. zur Beförd. d. ges. Naturw. zu Marburg, 1918, S. 26) geprägten Begriff der „pseudospontanen Progression" anwenden, wenn man auch annehmen wollte, daß, um mit W. zu sprechen, hier ein Charakter gleichsam in voreiliger Entwicklung auftritt, der normal und vollständig erst wesentlich später zur Entwicklung kommt. Der Name „pseudospontane Progression" ist nicht sinngemäß, den Begriff möchte ich dahin präzisieren, ein in der Entwicklungsrichtung liegender endogener Charakter wird im Verlaufe der geologischen Zeiten vorzeitig, nicht erblich, sondern als „spontane Pseudoprogression" ausgelöst. Wenn W. in diesem Zusammenhange auf das Altklugsein der Kinder hinweist, so möchte ich diese Erscheinung als etwas ganz anderes deuten, nämlich daß ein Charakter, der normal in der Entwicklung eines Individuums erst später in Erscheinung tritt, sich als Monstrosität, nicht erblich, zu früh einstellt. Durch zwei Beispiele erläutert:

1) spontane Pseudoprogressionen: unter den *Psiloceraten* des untersten Lias zeigen sich ganz vereinzelte Monstrositäten mit den typischen Rippenverdickungen der *Aegoceraten* des höheren unteren Lias; das gleiche ist bei einzelnen *Schlotheimien* der Fall. (Dr. Lange-Berlin, teilt mir gütigst mit, daß er unter zahlreichem normalen Material von *Psiloc. Johnstoni* zwei Exemplare mit *Aego-*

ceraten-Rippen fand; die gleiche Erscheinung fand er bei *Schlotheimia angulata*, *germanica* n. sp. LANGE und *angulosa* n. sp. LANGE. (Ich werde a. O. auf diese Erscheinungen zurückkommen.)

2) Es ist eine leicht zu beobachtende Tatsache, daß z. B. Rippen oder Knoten bei einzelnen Individuen bedeutend früher auftreten als bei der Mehrzahl der Formen.

Das Erscheinen der Zerschlitzung des Innenlobus bei *Tragophylloceras* könnte man vielleicht als eine Exzessivbildung auffassen, obgleich dies kein Charakter ist, der nach unseren Erfahrungen, zum Untergange führen muß.

Die Ableitung der *Stephanoceratida*.

Von NEUMAYR, POMPECKJ u. a. ist betont worden, daß alle *Ammonoideen* des Jura und der Kreide auf *Phylloceraten* der Trias zurückgeführt werden müßten, oder mit anderen Worten, daß die drei Stämme, *Stephanoceratida*, *Lytoceratida* und *Harpoceratida*, letzten Endes sich aus dem Stamm der *Phylloceratida* entwickelt hätten.

Der wesentliche Unterschied zwischen *Psiloceras calliphyllum* NEUM. (Tafel I, Fig. 11) gegenüber *Phylloceras* oder *Rhacophyllites* besteht nur darin, daß sich eine vollständige Zerschlitzung der Seitenwände des Innenlobus eingestellt hat, während der erste Innensattel noch monophyll bleibt. Im übrigen ist die Zerschlitzung der Lobenlinie mehr microphyll, als dies sonst bei jurassischen *Phylloceraten* der Fall ist. Die Sattelendigungen der Außenseite sind tetraphyll wie bei einigen *Phylloceraten*-Formen. Der Innenlobus der Gruppe des *Psil. laeve* QU. zeigt dem gegenüber das gleiche Bild wie die Gruppe des *Phyll. ibex* QU. (*Tragophylloceras*). Siehe Tafel I, Fig. 10.

Die Verschiedenheit in der Ausgestaltung der Lobenlinie bei den zu *Psiloceras* und *Schlotheimia* vereinigten Formen ist so groß, daß ich HYATT recht geben muß, wenn er (leider nur teilweise) diese Formen zu verschiedenen Gattungen auseinander zieht. Dagegen kann ich HYATT und anderen Forschern nicht darin folgen, die Gruppe des *Psil. planorbis* als eine in der Entwicklung der Lobenlinie gehemmte anzusehen. Dies könnte nur dann angenommen werden, wenn die Gruppe des *Psil. calliphyllum* die ältere wäre, während mit *Mojsvarites* zusammen die Gruppe des *Psil. planorbis* die ältere ist. Ich sehe mich daher auch veranlaßt anzunehmen, daß die Gruppe des *Psil. calliphyllum* (*Paraphylloceras gen. n.*) sich aus der Gruppe der *Mojsvariten* (vorbehaltlich des Studiums der Innenlobenlinie) an der Grenze Trias-Jura durch

Weiterbildung der Zerschlitzung der Seitenwände des Innenlobus und durch weitgehende microphyll-phylloceratenähnliche Zerschlitzung der Außenlobenlinie entwickelt hat. Daneben bleibt in der Gruppe des *Psil. laeve* der alte Charakter in der Ausbildung der Innenloben bestehen. Dieser Zweig allein dürfte sich in den Stamm der *Stephanoceratida* fortgesetzt haben, während *Paraphylloceras* mit gewissen *Schlotheimien* des alpinen Lias, Gruppe der *Schlotheimia marmorea* Opp., welche ähnliche Lobenlinien aufweist (vergl. Wähner, Beiträge z. Kenntnis des unteren Lias in den Alpen) auszusterben scheint. Die Furchenbildung auf der Außenseite der Schale und das winklige Zusammenstoßen der Rippen auf der Außenseite sind nur Excessivcharaktere, die nicht gestatten, stammeseinheitliche Gruppen auszuscheiden, selbst wenn die Formen in der gleichen Zone vorliegen. (Das gleiche gilt auch für die Kielbildung auf der Schalenaußenseite.) Sie müssen vielmehr in den nicht exzessiven Charakteren übereinstimmen, besonders neben der Berippung auch in dem Bauplane der Lobenlinien. Ohne über diesen Gegenstand eingehendere Studien angestellt zu haben, hat es keinen Wert für die Gruppe des *Schlotheimia marmorea* Oppel eine besondere Bezeichnung einzuführen.

Schlussfolgerungen.

Es ist hiermit festgestellt, daß sich in der oberen Trias aus monophyllen Formen (wie ich weiter oben gezeigt habe, aus solchen mit Lobenlinien vom Typ *Megaphyllites*) neben Weiterbildung anderer Charaktere der Lobenlinien durch Zerschlitzung der Seitenwände des Innenlobus der Stamm der *Stephanoceratida* herausgebildet hat. Also ein dritter Fall innerhalb der Ordnung der *Phylloceratida* (wenn wir diese auf die triadischen monophyllen Formen ausdehnen).

Und endlich kann ich noch einen vierten Fall bei *Rhacophyllites* anführen, wo, wie schon von Quenstedt gezeigt ist, bei *Rhacophyll. neojurensis* eine Seitenzacke im Innenlobus auftritt. Tafel I, Fig. 8.

Aus den vorhergehend erörterten Tatsachen läßt sich ableiten:

die gleiche Entwicklungsrichtung kann aus dem gleichen Stamme zu verschiedenen Zeiten unabhängig zum gleichen Resultat (Organisationshöhe) inbezug auf die Ausbildung eines bestimmten Charakters führen. Darüber hinaus läßt sich, wie ich im weiteren zeigen werde, ganz allgemein der Satz formulieren:

Die gleiche Entwicklungsrichtung führt in ver-

schiedenen Entwicklungsreihen zu gleichen oder verschiedenen Zeiten, oder in gleichen Entwicklungsreihen in verschiedenen Zeiten inbezug auf die Ausbildung bestimmter Charaktere zu dem gleichen Resultat (Organisationshöhe).

Die Art des Auftretens in der geologischen Zeitfolge zeigt, daß dies endogene Vorgänge sind.

Die Erforschung dieser Verhältnisse muß notgedrungen zur Diskussion der Frage nach der Vielstämmigkeit bestimmter bisher ausgeschiedener Gattungen, Familien und Ordnungen führen, eine Frage, die ich hier wohl anschneide, für deren Behandlung es aber noch nicht die Zeit ist.

Die Zwei- und Einspitzigkeit des Innenlobus.

Verfolgen wir die anderen Charaktere der Lobenlinie in ihrer Umänderung, so treten uns die gleichen Verhältnisse entgegen; auch in ihnen zeigt sich, daß die gleiche Entwicklungsrichtung in den verschiedenen Zeiten in den verschiedenen Reihen zum gleichen Resultat führt.

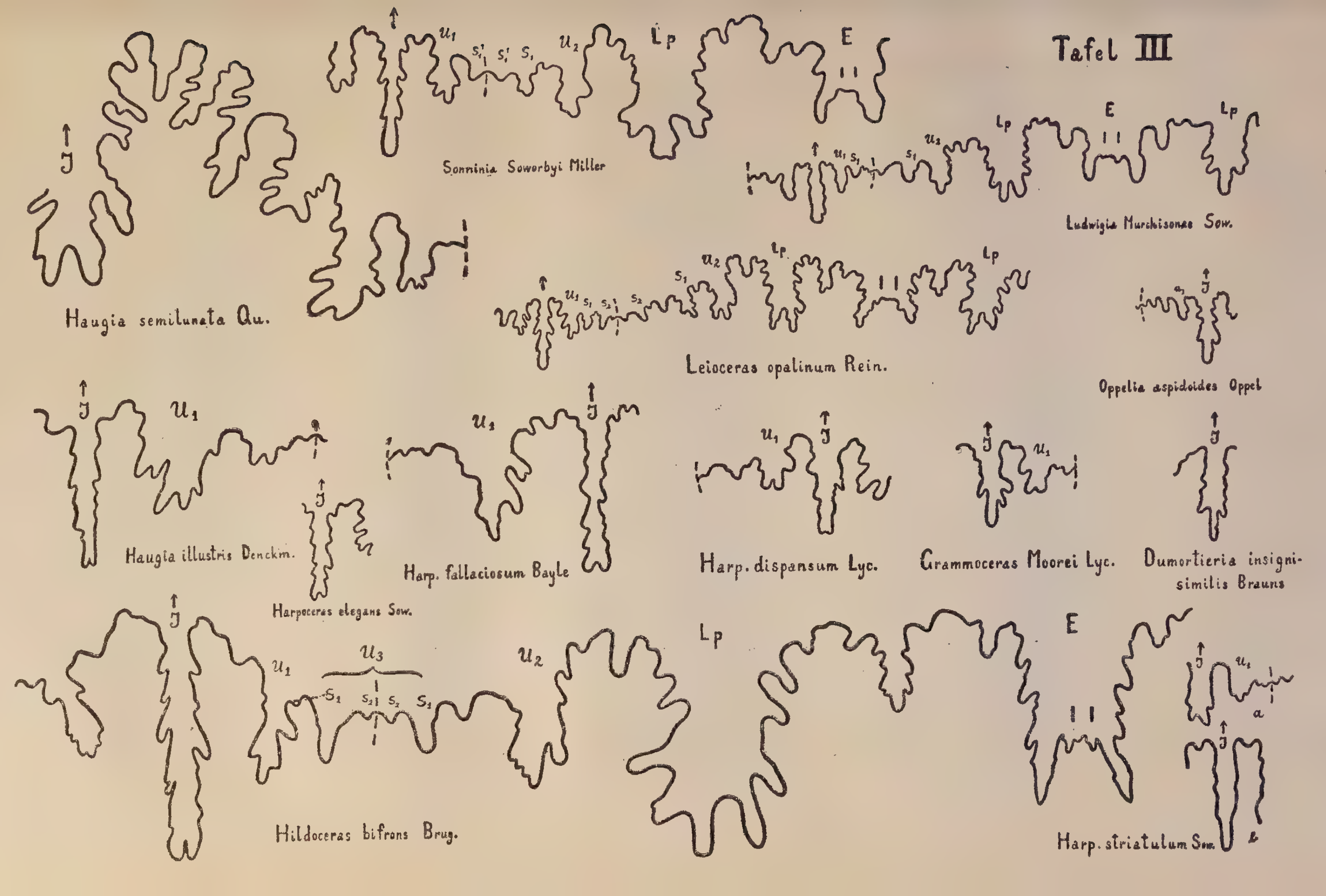
1. Der Innenlobus.

Die Zerschlitzung der Seitenwände des Innenlobus tritt uns bei allen Jura- und Kreide-Ammonoideen mit Ausnahme der reinen Reihen von *Phylloceras* entgegen. Die Wurzel der vier Jura-Kreide-Stämme haben wir schon in den monophyllen Formen vom *Megaphyllites*-Lobentyp gefunden, deren Innenlobus außer der Zweispitzigkeit keine weitere Zerschlitzung aufweist. Andererseits besitzen die triadischen *Arcestidae, Ptychitidae, Cladiscitidae* usw. einen Innenlobus mit zerschlitzten Seitenwänden (vergl. Tafel I, Fig. 1. *Gymnites Palmai*). Diese Formen gehören nicht zu den Vorläufern der Jura-Kreide-Ammonoideen, sodaß die Zerschlitzung des Innenlobus schon in der Trias unabhängig von den *Phylloceratida* und ihren Nachkommen wohl auch mehrfach aufgetreten ist.

a. Mediansattel im Innenlobus.

M in J ist bei den jurassischen Ammonoideen der Rückbildung unterworfen bis auf die beiden reinen Reihen *Phylloceras* und *Lytoceras*. Diese Rückbildung und der schließliche Übergang zur Einspitzigkeit (bezw. Dreispitzigkeit) wird von den *Stephanoceratida* im unteren Dogger erreicht. Die *Coeloceratinae* besitzen noch M in J, bei den *Stephanoceratinae* ist M in J nicht mehr ausgebildet (in ganz seltenen Fällen kann noch eine ganz leichte Andeutung

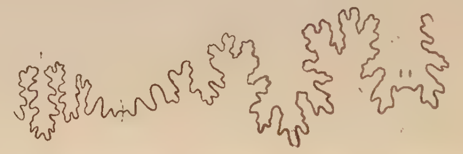

E
III
3a Tmetoceras scissum
1 Arietites rotiformis
3 Polymorphites polymorphum
2. Coel. crassum
6 Deroceras Davoei
Stephanoceras Humphriesi
7 Lytoceras fimbriatum
E
8 Parahoplites Milleti
Aspid. perar- matum
E
10 Leymeriella Schrammeni
E
uracanthites biformis
11
Douvilleiceras mammillare
E
Acanthoceras
Newboldi
12


Tafel III
Haugia semilunata Qu.
Sonninia Sowerbyi Miller
Ludwigia Murchisonae Sow.
Leioceras opalinum Rein.
Oppelia aspidoides Oppel
Haugia illustris Denckm.
Harpoceras elegans Sow.
Harp. fallaciosum Bayle
Harp. dispansum Lyc.
Grammoceras Moorei Lyc.
Dumortieria insignisimilis Brauns
Hildoceras bifrons Brug.
Harp. striatulum Sow.

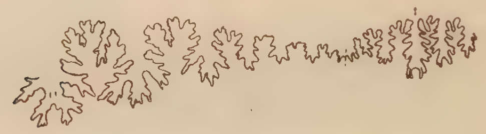
von zwei ganz kleinen Spitzen vorhanden sein). Bei den jüngeren *Stephanoceratinae* und den sich daraus entwickelnden Gruppen, wie *Perisphinctinae*, *Virgatitinae* (einschl. *Polyptichites*, *Astieria*, *Simbirskites*) und *Bidichotomitinae* (den sog. Tithon- und Neocom-*Hopliten* oder den kostaten *Hopliten* Uhligs) findet sich durchweg nur Einspitzigkeit, bezw. Dreispitzigkeit als Weiterbildung der Einspitzigkeit. Vergl. Tafel II, Fig. 1—6.

Die älteren *Harpoceratida* weisen mit *Polymorphites*, *Oxynoticeras*, *Amaltheus*, *Hildoceras*, *Harpoceras s. str.* e. p. M in J auf, dabei läßt sich die schrittweise Rückbildung des M verfolgen, bis schließlich bei einem Teil von *Harpoceras s. str.* und bei *Grammoceras* die Einspitzigkeit des Innenlobus erreicht ist. Etwas früher, wie bei den vorgenannten Gruppen, ist bei *Dumortieria* die Einspitzigkeit erreicht worden. (Vergl. Tafel III.) Also auch bei diesem Charakter wird sicher in zwei verschiedenen Stämmen, einmal im oberen Lias bei den *Harpoceratida* (und zwar bei den verschiedenen Gruppen nicht genau in der gleichen Zone), zum anderen im unteren Dogger bei den *Stephanoceratida* durch dieselbe Entwicklungsrichtung das gleiche Resultat erreicht.

Aus dem Dogger, Malm und der Unterkreide kenne ich, abgesehen wieder von den beiden Stämmen der *Phylloceratida* und *Lytoceratida*, nur Formen mit einspitzigem Innenlobus. Erst im Aptien-Albien tritt uns bei *Leymeriella* (e. p.), der Gruppe des *Desmoceras Hoyeri* Woll. (Textfig. 5), einigen Arten von *Puzosia* (Textfig. 6) und *Parahoplites* (Tafel II, Fig. 8) die Zweispitzigkeit



Figur 5. Desmoceras Keilhacki Wollem. Unt. Albien.



Figur 6. Puzosia latidorsata d'Orb.

des Innenlobus entgegen. Diese Gruppen können wir auf Grund ihrer sekundären Skulptur und der vorhandenen Übergänge wegen nur zu dem Stamm der *Lytoceratida* (Tafel II, Fig. 7—10) stellen. In ihrer Weiterentwicklung gehen die oben genannten Gruppen zu den verschiedenen Zeiten zur Einspitzigkeit über, so ist *Acan-*

thoceras (Tafel II, Fig. 12) durchweg ohne M in J, *Douvilleiceras* (Tafel II, Fig. 11) (als Gruppe der Formen, deren Rippen in Knotenreihen aufgelöst werden, ich verstehe nicht darunter das ungereimte in der vorhandenen Literatur) z. T. mit einspitzigem J. Im einzelnen müssen diese Verhältnisse an genau nach Zonen gesammeltem Material untersucht werden. In der oberen Kreide finden wir schließlich in dem Stamm der *Phylloceratida* wahrscheinlich einen mehrfachen Übergang von der Zweispitzigkeit des Innenlobus zur Einspitzigkeit bei einer Reihe von sog. Kreide-Ceratiten. Hierauf ist schon bei der Besprechung des Stammes der *Phylloceratida* hingewiesen. Auch hier müssen noch weitgehende Untersuchungen über die Hyattsche Monographie hinaus angestellt werden.

2. Die Ausbildung des Mediansattels im Außenlobus.

M in E scheint in seiner dachförmigen Ausbildung an das Vorhandensein von M in J gebunden zu sein. Beides sind alte Charaktere, welche den Formen aus dem Lias bezw. auch noch des unteren Dogger von ihren triadischen Vorläufern überkommen sind. Doch ist der Charakter in der dachförmigen Ausbildung von M in E nicht so prägnant wie das Vorhandensein von M in J. Ich schließe bei diesen Betrachtungen die *Phylloceratida* und *Lytoceratida* aus.

Der Übergang von der dachförmigen Form des M in J in die breitere rechteckige und meist niedrigere Form findet sich bei den *Stephanoceratida* und *Harpoceratida* an den Übergang von der Zweispitzigkeit zur Einspitzigkeit gebunden. Also auch dieser Charakter wird in verschiedenen Reihen unabhängig von einander zu verschiedenen Zeiten in gleicher Richtung mit dem gleichen Resultat umgestaltet.

4. Die sekundäre Zweispitzigkeit des Laterallobus und der Umschlagloben.

Als ein Hauptcharakter der Gattung *Lytoceras* gilt der mehr oder weniger symetrisch zweiteilige Bau besonders des primären Laterallobus. Die ontogenetische Entwicklung der Lobenlinie beweist aber, daß wir es primär mit einem triäniden Lobentyp zu tun haben, der sich erst sekundär im Laufe der Entwicklung des Individuums durch Herabdrängen der einen primären Lateralzacke und Rückbildung der primären Gegenzacke zu einem scheinbar primär zweizackigen oder dikraniden Lobentyp umwandelt. Man kann daher diese Ausbildung der Lobenlinie als pseudodikranid bezeichnen. (Vergl. Tafel II, Fig. 7 u. Fig. 12.)

Dieser Charakter wird zwar bei den *Lytoceraten* sehr zäh fest-

gehalten, auch bei vielen *Scaphiten*, *Turriliten* u. a. m. der oberen Kreide und gewissen *Acanthoceraten*, die ich auch auf Grund der Entwicklungsreihen zu den *Lytoceraten* stellen muß. Doch besitzen andere Formen, wie *Pleuracanthites*, *Holcodiscus*, *Puzosia*, *Leymeriella*, gewisse *Desmoceraten*, *Parahoplites* usw. keinen pseudodikraniden Lobentyp, obgleich ich sie doch auf Grund anderer Charaktere und der vorhandenen Entwicklungsreihen zu den *Lytoceratida* stellen muß. (Vergl. Tafel II, Fig. 8, 9, 10, 11.)

Andererseits werden die Lobenlinien einiger *Harpoceraten* des oberen Lias deutlich pseudodikranid. (Vergl. Tafel III, *Ludwigia* rechts oben.)

3. Die Lappung der Sättel.

Der Übergang der primären Dreilappigkeit des Sattels zur Zweilappigkeit mit einer primären Scheitelincision vollzieht sich in verschiedenen Entwicklungsreihen ebenfalls unabhängig von einander. Von Dietz (l. c.) ist dieser Werdegang bei den *Aegoceratinae* des unteren Lias gezeigt worden, er hat aber genau so unabhängig davon in dem Stamm der *Harpoceratida* stattgefunden und zwar erst an der Grenze vom Mittel- zum Oberlias. *Polymorphites*, *Oxynoticeras* und *Amaltheus* weisen dreilappige Sättel auf, während *Harpoceras*, *Hildoceras* und *Grammoceras* des oberen Lias nur zweilappige Sättel besitzen. Also auch dieser Charakter zeigt das gleiche Verhalten, daß die gleiche Entwicklungsrichtung in verschiedenen Reihen zu verschiedenen Zeiten zum gleichen Resultat führt.

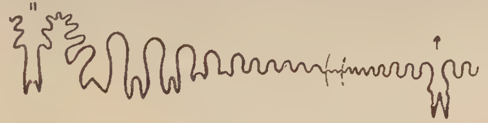
Es besteht übrigens ein fundamentaler Unterschied nicht nur zwischen der Entstehung der Zweilappigkeit bei *Rhacophyllites*, wo nur eine primäre Scheitelincision angelegt wird, gegenüber der Herausbildung der Zweilappigkeit bei den *Stephanoceratida* und *Harpoceratida*, sondern auch gegenüber *Phylloceras* ist der gleiche Unterschied festzustellen, worauf ich schon früher eingegangen bin.

Das erste Auftreten und die Bedeutung der Zweispitzigkeit des Innenlobus.

Im vorhergehenden habe ich darauf hingewiesen, daß die sämtlichen triadischen Ammonoideen, soweit wir davon Kenntnis besitzen, einen zweispitzigen Innenlobus haben. Dagegen kenne ich nach eigenen Beobachtungen und nach den Angaben der Literatur bei den Ammonoideen des Palaeozoikums nur das Vorkommen von einspitzigen Innenloben. Man hat das Auftreten des medianen

Sattels im Außenlobus mit der externen randlichen Lage des Sipho in Zusammenhang gebracht. Stimmt dies, so würde man das Auftreten des medianen Sättelchen im Innenlobus ebenfalls auf die gleiche Ursache zurückführen müssen, da nach Brancas Untersuchungen bei vielen Ammonoideen, besonders aus der Trias der Sipho in der frühsten Jugend intern randlich gelegen ist, und im Verlaufe der ontogenetischen Entwicklung nach dem Schalenrande der Außenseite wandert. Freilich weist die Lobenlinie in ihrer ontogenetischen Entwicklung noch in einem Stadium Einspitzigkeit des Innenlobus auf, wo der Sipho bereits zur Außenseite abwandert. Wir müßten also, die Abhängigkeit im obigen Sinne zugegeben, annehmen, daß durch die randliche interne Lage des Sipho in der frühen Jugend ein Einfluß auf die Entstehung des medianen Sättelchens im Innenlobus ausgeübt wird, dies aber erst in Erscheinung tritt, wenn der Sipho bereits zur Außenseite abwandert. Dadurch würde auch das Verschwinden der Zweispitzigkeit in geologisch späteren Zeiten seine Erklärung finden, daß schließlich die Einwirkung an Intensität abnimmt oder, wie aus den Untersuchungen Brancas zu vermuten ist, bei den Formen mit einspitzigem Innenlobus der Sipho in frühster Jugend nicht mehr eine intern randliche Lage besessen hat. Tatsächlich haben wir es mit einer langsamen Abnahme des M in J zu tun, bei der phylogenetischen Entwicklung. Diese Erörterungen erleiden dadurch keinerlei Änderungen, daß bisher in einem Fall in der ontogenetischen Entwicklung bei *Harpoceras striatulum* die Form von der Zwei- zur Einspitzigkeit übergeht. (Vergl. Tafel III, Figur rechts unten.)

Die Zweispitzigkeit wurde bei den Extrasiphonaten zeitlich zuerst bei *Medlicottia* aus dem Permo-Carbon beobachtet. Textfig. 7.



Figur 7. Medlicottia artiensis Gruen.

Originale im Göttinger Geol. Institut der Universität. Textfig. 1, 2, 3, 4, 5, 6, 7. Tafel I, Fig. 1, 2, 3a, 5, 6, 7, 9, 18, Tafel II, Fig. 2, 3, 3a, 8, 10, 11, 12 Tafel III sämtlich.

Tafel II, Fig. 3b in der Naturaliensammlung in Stuttgart. Alle übrigen Figuren nach der Literatur.

Zur Kenntnis der Petrographie des älteren Palaeozoicums zwischen Albungen und Witzenhausen, im Besondern des Variolits.

Von

O. Mügge.

Vorgelegt in der Sitzung vom 28. November 1919.

Zwischen dem Harz und dem rheinischen Schiefergebirge tritt älteres (vorpermisches) Palaeozoicum nur an 2 Stellen an die Oberfläche: im Werra-Tal zwischen Albungen und Witzenhausen und ca. 35 km südwestlich etwas westlich der Fulda bei Oberellenbach. Im Gebiet westlich der Fulda bestehen seine Schichten nach Beyschlag[1]) ausschließlich aus Grauwacken, im Gebiet an der Werra nehmen solche nach Moesta[2]) ebenfalls den weitaus größten Teil der Oberfläche ein, am Südrande dieses Gebietes erscheinen aber in den tiefsten Geländeteilen auch Tonschiefer mit Einlagerungen von Quarziten, Kieselschiefern, Hornsteinen, Kalken und Diabasen. Beyschlag betont für das Fulda-Gebiet die Ähnlichkeit der Grauwacken in petrographischer Hinsicht mit solchen des Kellerwaldes, bei dem Fehlen von Versteinerungen war aber eine sichere Altersbestimmung nicht möglich; ebenso wenig in dem viel ausgedehnteren Gebiet an der Werra, hier vergleicht Moesta die Grauwacken mit den Tanner Grauwacken und die Schiefer mit den Wieder Schiefern des Harzes.

Im Werra-Gebiet sind die Schiefer namentlich an den unteren Gehängen des Höllentals aufgeschlossen und enthalten hier

1) Beyschlag, Erläuterungen zu Bl. Altmorschen 1891.

2) Moesta, Bl. Witzenhausen und Bl. Allendorf (Erläuterungen von Beyschlag 1886).

zahlreiche Diabase eingeschaltet, welche nach Moesta auch endo- und exomorphe Kontakterscheinungen aufweisen. Die normalen Schiefer sind reich an Quarz, ziemlich serizitisch, etwas eisenschüssig und kohlehaltig, öfter sehr zierlich gefältelt und geknickt bis zur vollständigen Verruschelung. Ihre Kalkeinlagerungen sind dicht bis marmorartig, ein wenig flasrig, ohne bestimmbare organische Reste. Die namentlich westlich der Straße nach Apterode beobachteten Kieselschiefer sind voll von meist elliptisch deformirten Radiolarien, sie erscheinen auch im Kontakt mit den unten besprochenen Diabasen.

Die von Moesta als älter angesprochenen Grauwacken sind sandig, im Großen bankig, im Handstück fast kompakt. Obwohl durch Steinbrüche und neuerdings durch den Bahnbau im Gelstertal gut aufgeschlossen, wurden nur kümmerliche, nicht näher bestimmbare Reste von Muscheln, ? Crinoiden und ? Tentaculiten in ihnen gefunden. Ihre Mineralgemengteile Quarz, Orthoklos, Plagioklas, Biotit, Chlorit, Serizit u. a. weisen ebenso wie ihre, z. T. nur wenig abgerollten Gesteinsbruchstücke auf granitische und gneißige Gesteine und quarzige chloritische Serizitschiefer ähnlich etwa manchen Taunus-Schiefern als Ursprungsgesteine, die also heute nur in etwas größerer Entfernung anstehend bekannt sind; daneben aber auch, und zwar grade in den gröberen und fast noch scharfkantigen Bruchstücken auch auf spilitische und (selten) körnige Diabase, grobkörnige Kalksteine, Kieselschiefer und etwas kohlige Tonschiefer wie sie in den von Moesta als jünger angenommenen Schiefern vorkommen. Aus diesem Grunde und weil im Gelstertale die Grauwacke, nicht der Schiefer, vom Zechstein diskordant überlagert wird, weil ferner nur die Schiefer, nicht auch die Grauwacken, Diabase eingeschaltet enthalten, endlich auch weil Lagerungsverhältnisse welche auf jüugeres Alter der Schiefer hinweisen mir nicht bekannt geworden sind, möchte ich die Grauwacken für jünger als die Schiefer halten. Die Grauwacken könnten etwa, wie es Beyschlag für die petrographisch durchaus ähnlichen Grauwacken von Oberellenbach tut, zum Culm gestellt werden, wenn die Parallelisierung der Schiefer mit den Wieder Schiefern zu Recht besteht.

Die Diabase, welche in dem kleinen Schiefergebiet sehr zahlreich, in dem großen Grauwackengebiet nirgends auftreten, sind in den vorhandenen Aufschlüssen alle sehr zersetzt; indessen kann man an mehreren Stellen doch große Ähnlichkeit z. B. mit Harzer körnigen Diabasen feststellen, häufiger als dort scheint allerdings ein Olivingehalt zu sein. An der Grenze zu den Schiefern werden

sie zuerst feinkörnig, schließlich völlig dicht, an manchen Stellen offenbar glasreich, dabei öfter etwas, aber nicht sehr, blasig. In den dichten Gesteinen ist Olivin reichlich ausgeschieden, z. T. in größeren Einsprenglingen deren Durchschnitte in feine Spitzen auslaufen und welche manchmal voll von Grundmasse-Einschlüssen sind, dann aber namentlich in zahlreichen mikroskopisch-kleinen Einsprenglingen die im Längsschnitt als zweizinkige Doppelgabeln, im Querschnitt als abgestumpfte Rhomben mit großem Grundmasse-Einschluß erscheinen. Sie sind völlig zersetzt, sodaß nur aus diesen Formen noch auf ihre Olivin-Natur geschlossen werden kann. Zwischen ihnen bilden feine, zu besenförmigen Bündeln gruppierte Plagioklase mit einheitlicher Zwischenfüllung von Augit die reichlich vorhandene sehr fein-diabasisch struierte Grundmasse. Erz, Biotit und Apatit der körnigen Diabase fehlen hier.

Nur an der schon von Moesta genannten Stelle ganz nahe Bahnhof Albungen wurde die variolitische Varietät wieder aufgefunden. Ihre Variolen haben dieselbe Zusammensetzung wie die dichten Diabase sonst am Schieferkontakt, die Zwischenmasse der Variolen dagegen scheint Glas nur mit Ausscheidungen zahlreicher kleiner Olivine gewesen zu sein, es ist jetzt allerdings vollständig zu einem Aggregat von Carbonat, Chlorit und Serpentin zersetzt. Die Variolen stimmen auch in ihrer Struktur mit der Struktur der Grundmasse der dichten Diabase durchaus überein, namentlich auch lassen sie im Gegensatz zu den Sphaerolithen der sauren Ergußgesteine keinen Krystallisationsmittelpunkt (Ausgangspunkt) erkennen; daß sie sich trotzdem längs Kugeloberflächen von der umgebenden Glasmasse abgrenzen, dürfte in folgendem begründet sein. Vom jetzigen Mittelpunkte der Variolen in dem zuerst Feldspatkeime auftauchten, wuchsen diese anfangs strahlig nach allen Richtungen, wurden aber alsbald durch die schon ausgeschiedenen zahlreichen kleinen Olivineinsprenglinge, sehr bald auch durch die fast gleichzeitig einsetzende Krystallisation des Schmelzrestes zu Augit fortwährend unterbrochen; an jeder Unterbrechungsstelle entstand ein neues Wachtumszentrum, und nur weil diese Unterbrechungen wegen der großen Zahl, Kleinheit und regellosen Verteilung der Olivine auf allen Seiten gleichmäßig erfolgte, blieb die Grenze zwischen dem durch die Ausscheidung von Plagioklas und Augit völlig krystallin werdenden Teile des Magmas und jenem der nur aus Glas mit mikroskopischen Olivin-Einsprenglingen bestand zu jeder Zeit annähernd eine Kugelfläche. So kommt es, daß diese „Variolen“ den „Sphaerolithen“ der sauren Erguß-

gesteine, deren regelmäßig radialstrahliges Wachstum nicht durch das Vorhandensein zahlreicher kleiner älterer Einsprenglinge behindert wurde als kuglige Wachtumsformen von Faseraggregaten[1]) ohne regelmäßigen Bau gegenüberstehen.

Nach dieser Auffassung würden variolitische Grenzbildungen in Diabasen, welche nicht allzu häufig beobachtet sind, bedingt sein durch die frühzeitige Abscheidung der Olivine und rasche Abkühlung. Zu ersterem wird ein hoher Gehalt von Mg und Fe nötig sein, der aber die Alkalien und Tonerde nicht so weit zurückdrängen dürfte, daß es wie bei den Pikriten nicht mehr zur Ausscheidung von Plagioklas kommt. Die rasche Abkühlung scheint bei diesen Diabasen in der Tat vorhanden gewesen zu sein; die Art der Lagerung und die Veränderung der Struktur in der Nähe der Grenze sprechen durchaus für ein Ergußgestein und zwar vermutlich Ergüsse unter hohem Wasserdruck, was ja schnelle Abkühlung noch mehr als Erstarrung an der Erdoberfläche fördern und zugleich die geringe Gas-Entwicklung wie das Fehlen größerer Tuffmassen in der Nähe erklären würde.

Es stimmt damit auch das Fehlen irgend erheblicher exomorpher Kontakterscheinungen. Nirgend wurden im Kontakt spilositische oder desmositische Schiefer oder Adinolen beobachtet; auch an der von Moesta dafür angegebenen Stelle wurden nur sehr dichte, z. T. Kieselschiefern ähnliche Gesteine gefunden an denen, namentlich auf der verwitterten Oberfläche erscheinende Knötchen zwar Spilosit-ähnlich erscheinen können, in Wirklichkeit aber Durchschnitten von Radiolarien entsprechen.

Die im Laboratorium der Preußischen geologischen Landesanstalt ausgeführten chemischen Analysen der Variolite und ihrer Zwischenmasse ergeben größere Unterschiede zwischen beiden als nach der mikroskopischen Untersuchung erwartet wurde. Während die Variolen, entsprechend der Anreicherung an Plagioklas etwas höheren Gehalt an Alkalien (6.62% gegen 4.85%) aufweisen[2]), fällt in der Zwischenmasse der Variolite der große Gehalt an $MgO + FeO$ (23.86% gegen 12.60%) und der Mangel an Alkalien auf (2.16% gegen 6.62%). Letzteres dürfte daran liegen, daß der Schmelzrest tatsächlich stärker an Alkalien als an andern Bestandteilen verarmt war und daß die Alkalien, welche hier nur

1) Als Fasern im Querschnitt erscheinend; in Wirklichkeit also wohl Täfelchen.

2) Der Gehalt an CaO ist angesichts dessen leichter Beweglichkeit nicht von Bedeutung.

im Glas steckten, aus ihm viel mehr als aus den Variolen, wo sie im Plagioklas gebunden waren, entfernt wurden; die starke Anreicherung an $MgO + FeO$ dürfte ursprünglich sein; auch in frischen glasreichen Basalten beobachtet man, wenn zunächst Feldspate ausgeschieden sind an diesen vielfach dunkle, offenbar sehr eisenreiche „Bärte“ krystallinischer Ausscheidungen.

Materialien für eine wissenschaftliche Biographie von Gauss.

Gesammelt von **F. Klein, M. Brendel** und **L. Schlesinger.**

VII. Über die astronomischen Arbeiten von Gauss.

Von

M. Brendel in Frankfurt a. M.

Vorgelegt in der Sitzung vom 8. März 1918 von F. Klein.

Erster Abschnitt.

Theoretische Astronomie.

Einleitendes.

Wie auf allen Gebieten seines wissenschaftlichen Schaffens, so zeigt Gauß auch auf dem Gebiete der Astronomie das Streben, „seinen Untersuchungen die Form vollendeter Kunstwerke zu geben“ und „nie eine Arbeit zu veröffentlichen, bevor diese eine durchaus vollendete Form erhalten hatte“ [1]). Von den zahlreichen von ihm geplanten grundlegenden und den Gegenstand erschöpfenden astronomischen Werken hat er deshalb nur eines, die „Theoria motus corporum coelestium in sectionibus conicis solem ambientium“, Hamburg 1809, seinen Zeitgenossen geschenkt. In ihm behandelt er, soweit es die Bestimmung einer Bahn aus den Beobachtungen betrifft, ausschließlich die elliptische und die hyperbolische Bewegung. Vielleicht mag er mit Rücksicht darauf, daß Olbers die Aufgabe der parabolischen Bahnbestimmung mit Erfolg gelöst hatte, die weitere Durchführung seiner Untersuchungen über diesen Gegen-

1) Sartorius v. Waltershausen, Gauß zum Gedächtnis, Leipzig 1856, S. 82.

stand einstweilen zurückgestellt haben. Jedenfalls hat er hierüber nur die Abhandlung „Observationes cometae secundi A. 1813 in observatorio Gottingensi factae, adiectis nonnullis adnotationibus circa calculum orbitarum parabolicarum“, Göttingen 1813 (Werke VI, S. 25) veröffentlicht, in der er sich ausdrücklich auf das Olberssche Werk bezieht. Und doch beabsichtigte er, auch hier ein großes grundlegendes Werk zu schaffen, wie unter anderem aus der von Lindenau im November 1815 verfaßten Einleitung zu der Zeitschrift für Astronomie, herausgegeben von Lindenau und Bohnenberger, Erster Band, hervorgeht, wo es, S. 46, heißt: „Gauß hat neuerlich beinahe die ganze Kometentheorie umgearbeitet, neue Tafeln entworfen, und wir haben ein eigentümliches Werk darüber von ihm zu erwarten, was wohl nichts zu wünschen übrig lassen wird“. (Vgl. die Bemerkungen, Werke VII, 1906, S. 373—374.) Die Vorarbeiten für dieses Werk, die sich im Nachlaß vorgefunden haben, sind Werke VII, 1906, S. 323—373, zusammengestellt.

Ebenso hat die gewaltige Arbeit, die Gauß bei der Berechnung der Störungen der Pallas geleistet hat, erst bei der Veröffentlichung des VII. Bandes seiner Werke im Jahre 1906 der astronomischen Welt zugänglich gemacht werden können. Trotz sorgfältiger Durchsicht und Bearbeitung des Nachlasses wird natürlich nicht erwartet werden können, daß die ganze Fülle von Gedanken und Kunstgriffen, die Gauß hier zur Anwendung brachte, damit ans Tageslicht gezogen ist, umsomehr als außer dem Text der im VII. Band, S. 439 f. abgedruckten „Exposition d'une nouvelle méthode de calculer les perturbations planétaires“ der gesamte auf Pallas bezügliche Nachlaß fast nur aus einer nahezu unübersehbaren Zahl von Blättern mit Rechnungen ohne jeden erläuternden Text besteht. Die eigentliche Berechnung der Störungen, die bis auf einen kleinen Rest der Marsstörungen von Gauß ganz durchgeführt worden ist, hat sich zwar vollständig aufklären und veröffentlichen lassen, aber über die tiefer liegenden Überlegungen von Gauß und über die Gründe, warum er seine Entwickelungen überall gerade in der vorliegenden Art und Weise gemacht hat, sowie über andere Einzelheiten lassen sich mit Sicherheit keine Schlüsse ziehen. Was man darüber vermuten kann, wird weiter unten bei der eingehenden Besprechung dieser Arbeiten gesagt werden.

Dem gleichen Schicksal sind außer vielen anderen kleineren oder ganz unvollendeten Arbeiten auch Gauß' umfangreiche Untersuchungen über die Bewegung des Sonnensystems im Raume verfallen.

Daher kann man nur an der Hand des Nachlasses, nicht aber aus den gedruckten Abhandlungen ein Bild von der Entwicklung

der Gaußschen Arbeiten, besonders auf dem Gebiete der theoretischen Astronomie, gewinnen.

Der Beginn der Periode von Gauß' eingehendster Beschäftigung mit astronomischen Untersuchungen wird durch die Entdeckung des Planeten Ceres im Jahre 1801 eingeleitet. Doch hatte er sich schon vorher vielfach mit kleineren oder größeren astronomischen Problemen beschäftigt; über seine frühzeitig angestellten astronomischen Beobachtungen wird im zweiten Abschnitt dieses Aufsatzes „Praktische und sphärische Astronomie" berichtet werden.

In der Scheda Aa vom Jahre 1798, die mit der „Scheda prima de curva lemniscata" beginnt, steht auf S. 16 die Überschrift „Scheda secunda de motu cometarum". Darauf folgen aber die Differentialgleichungen des Zweikörperproblems mit Bezeichnungen, die sich auf die Erde beziehen und aus denen Gauß das Flächenintegral und einige weitere Beziehungen ableitet, ohne auf die Theorie der Kometenbahnen einzugehen. Da er aber in der Tagebuchnotiz Nr. 94 vom Juli 1798 sagt „Cometarum theoriam perfectiorem reddidi", so mag er in dieser Zeit einige weitergehende Entwickelungen über die parabolische Bewegung gemacht haben, die er dann nicht weiter ausgebaut hat und über die keine Aufzeichnungen erhalten sind. In der Scheda Ac sehen wir ihn Ende 1799 oder Anfang 1800 mit dem Studium von Olbers' Abhandlung über die Bahnbestimmung der Kometen[1]) beschäftigt, und in der Scheda Ae vom Sommer 1800 findet man mehrere ganz vereinzelte Notizen, die wohl darauf hinweisen, daß er sich dem Studium einzelner Werke der theoretischen Astronomie, im besonderen auch der Mechanik des Himmels hingab. Während seiner Studienzeit in Göttingen 1795—1798 hatte er nach Ausweis der Verzeichnisse der Göttinger Universitäts-Bibliothek von dieser u. a. folgende astronomische Werke entliehen: Lalande, Astronomie; Séjour, Traité analytique des mouvemens célestes; Cousin, Introduction à l'astronomie physique; Clairaut, Théorie de la Lune; Lagrange, Mécanique analytique; Newton, Opera omnia [2]).

I. Theorie des Mondes.

Das erste Unternehmen einer größeren Arbeit auf astronomischem Gebiete ist die Bearbeitung der Theorie des Mondes, die

1) Über die leichteste und bequemste Methode, die Bahn eines Kometen zu berechnen, Weimar 1797 (Wilhelm Olbers, Sein Leben und seine Werke, Band I, herausgegeben von C. Schilling, Berlin 1894).

2) Die Prinzipien Newtons hatte er schon 1794 erworben.

sich im Nachlaß auf zusammengehefteten Blättern befindet, in den Werken Band VII, 1906, S. 613 f. abgedruckt ist und nach der Tagebuchnotiz Nr. 120 („Theoriam motus Lunae aggressi sumus“) aus dem August 1801 stammt.

Bis zum Jahre 1788 waren zur Berechnung der Ephemeriden des Nautical Almanac die Mayerschen Mondtafeln[1]) im Gebrauch. Indessen war Mason vom Board of Longitude beauftragt worden, sie weiter zu verbessern, und seit 1789 traten daher die Masonschen Tafeln[2]) an die Stelle der ursprünglichen Mayerschen. Da sich bald das Bedürfnis einer weiteren Verbesserung der Tafeln fühlbar machte, so stellte die Pariser Akademie im Jahre 1798 die Preisaufgabe[3]): „Aus einer großen Anzahl der besten, zuverlässigsten alten und neuen Mondbeobachtungen, wenigstens 500 an der Zahl, die Epochen der mittleren Länge des Apogäums und des aufsteigenden Knotens der Mondbahn zu bestimmen.“

Bürg bearbeitete[3]) diese Aufgabe, indem er mehr als 3000 Beobachtungen benutzte, die er mit den Mayerschen Mondtafeln verglich, und indem er auch sonst über den Rahmen der gestellten Preisaufgabe hinausging. Er erhielt, ebenso wie Bouvard, der auch eine Abhandlung eingereicht hatte, den Preis im Jahre 1800. Bürg setzte seine Untersuchungen über die Mondbewegung fort, und auch Laplace begann um diese Zeit, seine Mondtheorie zu schaffen. Im Jahre 1800 setzte das Pariser Bureau des Longitudes einen neuen Preis[4]) aus für die Erfüllung folgender Bedingungen:

„1) Aus der Vergleichung einer großen Anzahl guter Beobachtungen den Wert der Koeffizienten der Mondungleichheiten auf das genaueste zu bestimmen und für die Länge, für die Breite und für die Parallaxe dieses Gestirns genauere und vollständigere For-

---

1) Tabulae motuum Solis et Lunae novae et correctae, London 1770. — T. Mayer hatte zuerst seine „Novae tabulae motuum Solis et Lunae“ in den Comm. scient. Gott., Göttingen, 1752 veröffentlicht, diese sodann weiter verbessert und die verbesserten Tafeln 1755 zum Wettbewerb um den für die Längenbestimmung vom Board of Longitude ausgesetzten Preis nach London geschickt. Die Entscheidung über den Preis zog sich in die Länge; T. Mayer starb 1762 und seine Witwe sandte seine Handschriften, die eine weitere Verbesserung der Tafeln enthielten, ebenfalls nach London. Dort wurde seine „Theoria Lunae iuxta systema Newtonianum“, 1767 und die eigentlichen Tafeln „Tabulae motuum Solis et Lunae novae et correctae“ 1770 gedruckt. Die Witwe erhielt einen Teil des Preises. — Vergl. R. Wolf, Geschichte der Astronomie, München 1877, S. 497.

2) Charles Mason, Tob. Mayers Lunar tables improved, London 1787. — Vergl. auch Lalande, Astronomie II, 3. Ausgabe, Paris 1792, S. 175.

3) Monatl. Corr. 1800, Mai, S. 541.

4) Monatl. Corr. 1800, August, S. 165.

meln zu geben als diejenigen sind, auf welchen die bisher gebrauchten Mondtafeln beruhen.

2) Aus diesen Formeln Mondtafeln mit einer hinlänglichen Bequemlichkeit und Sicherheit für die Berechnung zu entwerfen."

Die Erteilung des Preises war an keinen Zeitpunkt gebunden.

Vielleicht ist hierin der Anlaß zu suchen, daß Gauß die Bearbeitung der Mondtafeln in Angriff nahm. Er leitete als Fundamentalgleichungen die Differentialgleichungen des reziproken kurtierten Radiusvektor, der mittleren Länge (oder der Zeit) und der Tangente der Breite ab und benutzte die wahre Länge als unabhängige Veränderliche. Seine Fundamentalgleichungen sind also ähnlich denen von Clairaut[1]), d'Alembert[2]) und den später von Laplace[3]), Plana[4]) u. a. aufgestellten, nicht aber den Eulerschen[5]).

Die Integration wird durch Annäherungen ausgeführt, bei denen die Entwickelungen nach Potenzen der Exzentrizität und der Tangente der Neigung fortschreiten. Die Integrationsdivisoren entwickelt Gauß nach Potenzen des Verhältnisses der mittleren Bewegungen von Mond und Erde. Die Form, die die Resultate dadurch erhalten, stimmt also im wesentlichen mit denen der späteren Planaschen Theorie[4]) überein. Das Ergebnis der ersten Annäherung vergleicht Gauß mit den Werten von Tobias Mayer. Indessen hat er die ganze Arbeit bald wieder aufgegeben und nur die Störungen der Breite durchgeführt; sie ist offenbar plötzlich abgebrochen worden, wofür sich die Erklärung aus der Einleitung zur Theoria motus und aus einem Briefe an Schumacher vom 23. Januar 1842 ergibt, in dem es heißt: „Eben im Sommer 1801 hatte ich mir vorgesetzt, ähnliche Arbeit über den ☽ [Mond] auszuführen; aber kaum hatte ich die theoretischen Vorarbeiten angefangen (denn diese sind es, auf welche in der Vorrede meiner Th[eoria] M[otus] C[orporum] C[oelestium] angespielt wird), als das Bekanntwerden von Piazzis ⚳ [Ceres] Beob[achtungen] mich in eine ganz andere Richtung zog". Die in diesem Briefe erwähnte Stelle aus der Vorrede zur Theoria motus sehe man im Abdruck des Tagebuchs bei der Nr. 119, Werke X 1, S. 561. — Über den Fortschritt der

1) Clairaut, Théorie de la Lune, St. Petersburg 1752, 1765.

2) D'Alembert, Recherches sur différents points importants du système du monde, Paris I. Band 1754, III. Band 1756.

3) Laplace, Mécanique céleste, Tome III., Livre VII, Paris 1802.

4) Plana, Théorie du mouvement de la Lune, Turin 1832.

5) Euler, Opuscula varii argumenti, Berlin 1746; Theoria motus Lunae exhibens omnes eius inaequalitates, Petersburg 1753 und Theoria motuum Lunae nova methodo pertractata, Petersburg 1772.

Untersuchungen von Laplace und Bürg finden sich ausführliche Berichte in der Monatl. Corr. 1800—1802 und dies mag dazu beigetragen haben, daß Gauß auch späterhin seine Arbeit anscheinend nicht wieder zur Hand genommen hat. Laplace's Ergebnisse erschienen 1802 im dritten Bande der Mécanique céleste; Bürg erhielt 1803 den neuen Preis, während der Druck seiner Mondtafeln[1]) sich bis zum Jahre 1806 hinzog. Im Jahre 1803 hat Gauß in der Scheda Am die Ergebnisse von Laplaces Untersuchungen aus der Mécanique céleste (Band III) herausgeschrieben und einige Notizen dazu gemacht, die sich auf die Tafeln von Mason und auf die bis dahin bekannt gewordenen Ergebnisse von Bürgs Untersuchungen zu beziehen scheinen.

II. Bahnbestimmung aus drei oder vier Beobachtungen (im besonderen Theoria motus).

1. Die Entdeckung der Ceres.

Die von Piazzi in Palermo am 1. Januar 1801 gemachte Entdeckung der Ceres wurde erst im Mai durch die deutschen Zeitungen bekannt und die ersten genaueren Nachrichten darüber gab das Juni-Heft der „Monatlichen Correspondenz zur Beförderung der Erd- und Himmelkunde", die v. Zach, Oberstleutnant und Direktor der Sternwarte Seeberg bei Gotha, seit dem Jahre 1799 herausgab und die als Sammelpunkt wichtiger neuer geographischer und astronomischer Nachrichten diente. Piazzi hatte am 24. Januar einen Brief an Professor Bode, Direktor der Berliner Sternwarte, und zu gleicher Zeit an Oriani in Mailand und auch an Lalande in Paris gerichtet, in dem er mitteilt, daß er einen sehr kleinen Kometen ohne Nebel und Schweif im Sternbilde des Stiers entdeckt habe[2]). Im Februar berichtete Lalande darüber zwar an v. Zach, ohne aber den näheren Ort am Himmel anzugeben, sodaß v. Zach auf weitere Nachricht wartete. Die Briefe von Piazzi an Oriani und Bode gelangten erst im April an ihre Bestimmungsorte — der an Bode war 71 Tage unterwegs —, während Piazzi das Objekt nur bis zum 11. Februar hatte verfolgen können. In seinen Briefen gab Piazzi nur zwei beobachtete Örter vom 1. und 23. Januar, roh auf ganze Mi-

1) J. T. Bürg, Tables de la Lune in den Tables astronomiques publiées par le Bureau des Longitudes de France, Première Partie, à Paris, chez Courcier, 1806.
2) Monatl. Corr. 1801, Juni, S. 604.

nuten abgerundet, und bemerkte nur noch, daß vom 10. auf den 11. Januar die rückläufige Bewegung in rechtläufige übergegangen sei; im übrigen fügte er in seinem Briefe an Oriani hinzu, er habe die Vermutung, daß es sich um einen Planeten handele, während er in den Briefen an Bode und Lalande nur von einem Kometen spricht.

Beide, Bode sowohl wie Oriani, gaben die Nachricht sogleich an v. Zach weiter, der dann in dem erwähnten Hefte der Monatlichen Correspondenz einen ausführlichen Artikel „Über einen zwischen Mars und Jupiter längst vermuteten, nun wahrscheinlich entdeckten neuen Hauptplaneten unseres Sonnensystems“ brachte, während Bode die Entdeckung der Kgl. Preußischen Akademie der Wissenschaften mitteilte und für die Bekanntmachung in einigen Zeitungen sorgte.

v. Zach hatte gerade im September 1800 bei Gelegenheit einer „kleinen astronomischen Reise nach Celle, Bremen und Lilienthal“, über die er in den Heften der Monatlichen Correspondenz der Jahre 1800 bis 1801 ein ausführliches Tagebuch veröffentlichte, mit fünf anderen Astronomen (Schröder, Harding, Olbers und wahrscheinlich v. Ende und Gildemeister), die sich in Lilienthal trafen, beschlossen, „eine geschlossene Gesellschaft von 24 praktischen in ganz Europa verbreiteten Astronomen“ zu gründen[1]), die sich das Aufsuchen des zwischen Mars und Jupiter vermuteten Planeten durch gleichzeitige Verbesserung der Sternverzeichnisse angelegen sein lassen sollten. Auch Piazzi befand sich unter den 24 ausgewählten Astronomen, hatte aber die Einladung, an der Gesellschaft teilzunehmen, noch nicht erhalten[2]). Bode sowohl wie Oriani hielten daher gleich an der Auffassung fest, daß das neue Objekt ein Planet sei, der sich zwischen Mars und Jupiter bewegt, und auch v. Zach trat natürlich dieser Auffassung bei, umsomehr als die oberflächliche Berechnung einer Kreisbahn darauf hindeutete. v. Zach versuchte eine etwas schärfere Berechnung einer Kreisbahn, die eine merkwürdige Ähnlichkeit mit der Bahn des Kometen von 1770 zeigte[3]); er fragte sich, ob beide Objekte nicht vielleicht identisch sind; gewisse Zweifel an der Natur des neuen Planeten bestanden also immer noch: Piazzi spricht in einem späteren Briefe von ihm als von einem Kometen und auch die Pariser Astronomen scheinen

1) Monatl. Corr. 1801, Juni, S. 602.

2) A. a. O., S. 603.

3) A. a. O., S. 614; es handelt sich um den Kometen 1770 I (Nr. 102 des Galleschen Verzeichnisses); die halbe große Axe ist das einzige Element, das bei beiden Bahnen Ähnlichkeit zeigt.

der gleichen Ansicht gewesen zu sein und daher weniger Wert auf die Entdeckung gelegt zu haben[1]).

Bode hatte inzwischen an Piazzi mit der Bitte um genaue Mitteilung seiner Beobachtungen geschrieben, erhielt aber zunächst keine befriedigende Antwort. Es entspann sich ein eingehender Briefwechsel über Piazzis Entdeckung zwischen Bode, v. Zach, Olbers, der die Nachricht aus der Zeitung erfahren hatte, und Burckhardt in Paris, wohin Piazzi endlich in einem weiteren Briefe an Lalande, der in Paris am 31. Mai eintraf, seine Beobachtungen genauer mitgeteilt hatte, aber mit der Bitte, sie vorläufig nicht zu veröffentlichen; den deutschen Astronomen teilte sie Burckhardt unter der gleichen Bedingung mit. Auf Grund der genaueren Beobachtungen berechnete Burckhardt auch schon eine Ellipse; seine Versuche, die Beobachtungen durch eine Parabel darzustellen, scheiterten. Die entsprechenden Elemente sind im Juli-Heft der Monatlichen Correspondenz mitgeteilt, in der v. Zach in allen Monatsheften „Fortgesetzte Nachrichten über einen ..... neuen Hauptplaneten“ gab.

Die vollständigen Beobachtungen von Piazzi vom 1. Januar bis 11. Februar 1801 wurden endlich im September-Hefte der Monatl. Correspondenz veröffentlicht, nachdem sie Piazzi mit einigen Verbesserungen an Bode, Lalande und Oriani geschickt hatte, und gelangten so auch in die Hände von Gauß.

Im Oktober-Heft der Monatl. Corr. sagt v. Zach, daß gegen Mitte August bis Ende September nun von fast allen Astronomen Versuche gemacht wurden, den aus den Strahlen der Sonne wieder austretenden Planeten aufzufinden, aber ohne Erfolg; es herrschte auch allgemein in dieser Zeit schlechte Witterung. Die von Burckhardt berechnete elliptische Bahn war unsicher, nicht so sehr deswegen, weil das beobachtete Bahnstück ziemlich klein war, was damals von den Astronomen als die Hauptschwierigkeit[2]) empfunden wurde, sondern deswegen, weil er von einer willkürlichen Annahme über die Lage des Perihels ausging. Die Aufgabe, eine noch völlig unbekannte Planetenbahn aus den Beobachtungen zu bestimmen, war bisher nur beim Uranus aufgetreten, und hier konnte man zunächst eine Kreisbahn rechnen und sodann mit Hilfe entfernt liegender früherer Beobachtungen von Flamsteed 1690 und Tobias Mayer 1756, die Bode auffand, die Bahn genauer bestimmen. Olbers, der ebenfalls mit wenig Aussicht auf Erfolg die Be-

---

1) Monatl. Corr. 1801, Juli, S. 56—57.

2) Vgl. Theoria motus, Einleitung (Werke VII, S. 7).

rechnung einer elliptischen Bahn begonnen hatte, empfiehlt, der Vorausberechnung eine Kreisbahn zugrunde zu legen, deren Elemente er auch angibt; denn: „War der neue Planet vor dem 1. Januar durch sein Aphelium gegangen, so vermehrt sich seine heliozentrische Geschwindigkeit immer, und auch seine geozentrischen Längen müssen im August und September größer sein, als nach der Kreis-Hypothese. Ist er aber im Februar durch sein Perihelium gegangen, so hat sich nachmals die heliozentrische Geschwindigkeit vermindert und seine geozentrischen Längen müssen im August und September kleiner sein, als nach der Kreis-Hypothese. Weil man nun nicht wissen kann, welcher von beiden Fällen eintritt, so ist es zur künftigen Aufsuchung des Gestirns sicherer, die aus der Kreis-Hypothese gefolgerten Örter zum Grunde zu legen, die von den wahren nicht sehr abweichen können, und die unter beiden möglichen Fällen das Mittel halten“ [1]). Olbers, wie auch Burckhardt nahmen fälschlich an, daß der Planet zur Zeit seiner Entdeckung nicht weit entweder vom Perihel oder vom Aphel gestanden, während Gauß später zeigte, daß er sich nahezu zwischen beiden befunden habe [2]).

Piazzi verfaßte eine kleine Schrift [3]), in der er über die erste Entdeckung und die weiteren Beobachtungen eingehende Mitteilungen machte und auch, außer einer von ihm selbst berechneten Kreisbahn, die ihm durch Oriani zugesandten, von den anderen Astronomen berechneten Bahnen angibt. Der Schrift ist auch ein nochmals verbessertes Verzeichnis seiner Beobachtungen beigefügt. Im November-Heft der Monatl. Corr. gibt v. Zach einen ausführlichen Auszug aus dieser Schrift, „welche wahrscheinlich nicht so leicht und nicht so bald in den deutschen Buchhandel kommen dürfte“, und druckt auch die verbesserten Beobachtungen ab, die außer kleineren Änderungen eine Verbesserung der Rektaszension des 11. Februar um 15″ gegenüber den im September-Heft abgedruckten Beobachtungen enthalten.

Mit Rücksicht auf Olbers' erwähnten Vorschlag, die Vorausberechnung zur Wiederauffindung des Planeten auf eine Kreisbahn zu gründen, berechnete v. Zach eine Ephemeride [4]) für November und Dezember, um „dadurch allen Astronomen und Liebhabern der

1) Monatl. Corr. 1801, Oktober, S. 367.

2) A. a. O., 1801, Dezember, S. 639 (Bd. VI, S. 201).

3) Risultati delle Osservazioni della nuova Stella scoperta il di' 1 Gennajo all' Osservatorio Reale di Palermo, Palermo 1801.

4) Monatl. Corr. 1901, November, S. 578—581.

Sternkunde, die sich mit der Aufsuchung des Gestirns beschäftigen wollen, einen kleinen Dienst zu erweisen".

Inzwischen hatte sich Gauß, der die Monatl. Corr. in Braunschweig erhielt[1]), im Stillen an die Arbeit gemacht; das Interesse für den neuen Planeten veranlaßte ihn, seine Mondtheorie und seine rein mathematischen Untersuchungen liegen zu lassen[2]). In seinem Tagebuch finden sich aus dem Jahre 1801 die Notizen Nr. 119: „Methodus nova simplicissima expeditissima elementa orbitarum corporum coelestium investigandi; — Sept. m[edio] Brunsvigae", und Nr. 120: „Formulas permultas novas in Astronomia Theorica utilissimas eruimus. — Oct.".

Die ältesten im Nachlaß (Schedae Ag, Ah, Handbuch Bb) vorhandenen Aufzeichnungen über die Ceres stammen erst aus dem Anfang November und über Gauß' Untersuchungen in den ersten Wochen läßt sich daher keine volle Klarheit gewinnen. Einen Niederschlag der in der Tagebuchnotiz Nr. 120 genannten „Formulae permultae" mag ein in Band XI 1 der Werke abzudruckendes Stück aus dem Handbuch Bb geben.

Man sieht aber soviel, daß Gauß sofort den Plan faßte, wirklich neue brauchbare Methoden zur Bahnbestimmung zu schaffen, und daß er dabei von dem Gedanken ausging, sich nicht auf Versuche nach der einen oder anderen Richtung und auf irgend welche hypothetischen Voraussetzungen zu beschränken, sondern systematisch eine Bahn zu finden, die sich so gut wie möglich an die Beobachtungen anschließt: wenn die Piazzischen Beobachtungen auch nur 41 Tage umfassen, so muß es doch eine Ellipse geben, die sie am besten darstellt und die geeignet ist, die zur Wiederauffindung vorausberechneten Örter möglichst nahe anzugeben. Es galt also, eine Ellipse zu finden, die von allen willkürlichen Voraussetzungen frei war, und auf den ersten Seiten des Handbuchs Bb vom November 1801 finden wir diese Aufgabe bereits vollständig gelöst, wenn auch in einer weniger vollkommenen Form, als in der Theoria motus.

In einer kleinen Handschrift „Summarische Übersicht der zur Bestimmung der Bahnen der beiden neuen Planeten angewandten Methoden" hat Gauß seine ältesten Methoden zusammengestellt,

1) Gauß' Exemplare der Monatl. Corr. aus den Jahren 1800—1813 befinden sich in der Gauß-Bibliothek und enthalten manche handschriftliche Eintragungen von seiner Hand.

2) Vgl. den Abdruck des Tagebuchs, Werke X 1, S. 561—563 und oben S. 5.

und diese Handschrift hat er am 6. August 1802 an Olbers geschickt[1]), von dem er sie im November 1805 zurückerhielt. Kurz nach dem Erscheinen der Theoria motus bekam sie v. Lindenau, vermutlich bei einem Besuch bei Gauß, zu Gesicht und veröffentlichte sie mit Gauß' Zustimmung in der Monatl. Corr. im September 1809 (Werke VI, S. 148).

2. Einleitendes über Gauß' Methoden der Bahnbestimmung.

Obwohl die Bestimmung der Bahn eines Himmelskörpers aus drei gegebenen Beobachtungen eine fest umschriebene Aufgabe ist, so ist doch bekanntlich eine explizite Lösung nicht durchführbar, weil die beobachteten Örter und die zu bestimmenden Elemente der Bahn in sehr verwickelten Beziehungen zu einander stehen. — Man ist auf die Lösung der Aufgabe durch Annäherungen angewiesen und daher ist die Aufstellung einer fast unbegrenzten Anzahl von Methoden möglich, die durch mehr oder minder durchgreifende Unterschiede gekennzeichnet sind.

Man wird erwarten, daß Gauß das Gebiet, auf dem sich diese Methoden bewegen können, nach allen Richtungen hin durchforscht hat; man wird aber auch verstehen, daß er dies nicht sofort bei seinen ersten Bahnberechnungen hat durchführen können, da hier die numerische Bearbeitung des Einzelfalls drängte, damit die Bahn des neuen Planeten so bald wie möglich mit ausreichender Genauigkeit zum Zweck der Wiederauffindung am Himmel bekannt wurde. So erklärt es sich, daß Gauß' erste Bahnberechnungen zwar auf einem wichtigen neuen Grundgedanken beruhen, dem die Willkürlichkeit älterer Methoden nicht mehr anhaftet, daß sie aber in der einzelnen Durchführung noch nicht die Vervollkommnung und Verfeinerung zeigen, wie die Methoden der Theoria motus.

Dieser erste grundlegende Gedanke besteht in der Aufstellung einer Gleichung zwischen den Abständen des Planeten von der Sonne und von der Erde in der mittleren Beobachtung; es ist dies die im Folgenden mit 3) bezeichnete Gleichung (S. 15). Sie steht auf den in Band XI 1 der Werke abzudruckenden ersten Seiten des Handbuchs Bb als Gleichung I und findet sich wieder als Gleichung (7) des Artikels 6 in der soeben erwähnten „Summarischen Übersicht“ (Werke VI, S. 159). Gauß sagt dort von ihr: „Diese Formel ... ist der wichtigste Teil der ganzen Methode und ihre erste Grundlage“

---

1) Vgl. die Briefe vom 6., 18. August, 11., 14., 21. Sept., 10. Okt. 1802, 29. Okt., 2. Nov. 1805 aus dem Briefwechsel zwischen Gauß und Olbers.

und schreibt in dem die Sendung begleitenden Briefe an Olbers (vom 6. August 1802): „Der wesentlichste Punkt derselben [d. i. der Methode] ist die Formel (7) im 6. Artikel, auf die ich vor beinahe einem Jahr auf einem ganz anderen Wege kam. Sie ist gewissermaßen das Pendant zu der Ihrigen[1)] und man könnte sagen, daß beide ein zweites Differential brauchen; die Ihrige braucht die Veränderung der geozentrischen Geschwindigkeit, um das Verhältnis der Abstände, die meinige braucht die Veränderung der Richtung, um die Abstände selbst zu finden. Daß es außer der Ihrigen Formel noch eine ähnliche geben müsse, hatte ich vor fünf Jahren geahnt, da ich zum erstenmal Ihre Bestimmung der Kometenbahn las; ich äußerte damals etwas darüber gegen den sel. Lichtenberg, der mich sehr aufmunterte, mich in die Untersuchung einzulassen, allein meine damaligen sehr eifrigen Beschäftigungen mit der höheren Arithmetik, sowie mit Untersuchungen aus einem anderen Fache der Analyse[2)], worüber ich Ihnen in Zukunft einmal schreibe, brachten mir den Gegenstand bald wieder aus dem Sinne. Als ich im vorigen Jahre ganz unvermutet auf die Formel geriet, sah ich sogleich, von welchem Werte sie zur Abkürzung der ersten Annäherungsversuche bei einer von Hypothesen unabhängig sein sollenden[3)] Bestimmung der Bahn eines Himmelskörpers sein müsse. Glücklicherweise erhielt ich um die Zeit gerade die Piazzischen Beobachtungen im September-Heft der Monatl. Corr., an denen ich mich sogleich eine Probe der Methode zu machen entschloß. Die Verschiedenheit meiner Resultate von den vorhergegangenen bestimmte mich, sie bekannt zu machen, und die ferneren Rechnungen, worin ich dadurch verwickelt wurde, veranlaßten die fernere Ausbildung der Methode“.

In dem Briefe an Olbers vom 25. Mai 1802 sagt Gauß, daß er auf seine Fundamentalformel auf bizarrem Wege gekommen sei. Vielleicht darf man hiernach vermuten, daß er zuerst die streng giltigen Gleichungen aufgestellt hat, die sich im Handbuch Bb neben den genähert giltigen finden und die in Band XI 1 der Werke abgedruckt werden sollen.

1) Nämlich der bekannten Olbersschen Beziehung zwischen den Abständen von der Erde im ersten und dritten Ort, Gleichung 6) auf S. 16.

2) Gemeint sind wohl die aus dem Jahre 1797 stammenden Untersuchungen über lemniskatische Funktionen; vgl. überhaupt die Nummern 50—82 des Tagebuchs, Werke X 1, S. 509—525, in denen über Gauß' mathematische Untersuchungen während des Jahres 1797 berichtet wird.

3) Diese Worte bestätigen die oben ausgesprochene Vermutung, daß Gauß sogleich daran dachte, eine von willkürlichen Voraussetzungen freie Methode der Bahnbestimmung zu finden.

Man wird wohl nicht fehlgehen in der Annahme, daß die Auffindung dieser Gleichung den Anlaß zur Tagebuchnotiz Nr. 119 „Methodus nova simplicissima expeditissima elementa orbitarum corporum coelestium investigandi. Sept. m[edio]. Brunsvigae“ gab, mit der auch die Zeitangabe im vorstehenden Briefe übereinstimmt.

Olbers richtete, nachdem er die „Summarische Übersicht“ gelesen, einige darauf bezügliche Fragen an Gauß und bemerkt im Briefe vom 11. Sept. 1802: „Übrigens wird Ihnen die große Analogie Ihrer Hauptgleichung (7) mit der La Placeschen (Mechanik des Himmels, übersetzt von Burckhardt, I. Teil, p. 279 oder Mém. de l'Acad. Roy. de Paris 1780, p. 31) bekannt sein“[1]). Gauß antwortet am 14. Sept.: „Die La Placesche Formel, die ich vor vielen Jahren in seiner Théorie du Mouv. Ellipt. gesehen hatte, war mir ganz aus dem Gedächtnisse gekommen, bis ich ganz vor kurzem die Mécanique Céleste erhielt. Mich dünkt, sie muß sich sehr leicht aus (7) herleiten lassen“.

In der Theoria motus finden wir die Gaußsche Hauptgleichung nicht mehr in der gleichen Form; er ist hier noch einen Schritt weiter gegangen, indem er die von ihm mit P und Q bezeichneten Größen einführt, von denen weiter unten gesprochen werden soll.

Auch sind die gesamten mehr auf Interpolation beruhenden ersten Gaußschen Rechnungsmethoden in der Theoria motus durch systematisch vorgehende ersetzt, die er erst in den Jahren 1805—1807 schuf; auf sie beziehen sich die Tagebuchnotizen Nr. 125, 126, 127, 129. Darüber sagt Gauß im Briefe an Olbers vom 3. Februar 1806: „Ich habe in diesem Jahre fleißig an meiner Methode, die Planetenbahnen zu bestimmen, gearbeitet; obgleich bisher nicht so sehr an der Ausarbeitung, als an der größeren Vervollkommnung einzelner Teile derselben. Manches, glaube ich, ist mir gut gelungen und hat wenigstens eine von der vorigen ganz verschiedene Gestalt bekommen“, und in der Vorrede zur Theoria motus selbst: „Methodi enim ab initio adhibitae identidem tot tantasque mutationes passae sunt, ut inter modum, quo olim orbita Cereris calculata est, institutionemque in hoc opere traditam vix ullum similitudinis vestigium remanserit“ (Werke VII, 1906, S. 8—9).

Indessen hat Gauß gelegentlich den einen oder anderen Teil seiner älteren Methoden auch in der Theoria motus erwähnt (Vgl. die Einleitung, Werke VII, 1906, S. 9).

---

1) Laplace, Mécanique céleste, Première Partie, Livre I, No. 31 (Band I, S. 207 der Originalausgabe von 1799).

Wie Gauß auch in art. 2 der „Summarischen Übersicht" hervorhebt, stützt sich die erste Berechnung der völlig unbekannten Bahn eines Himmelskörpers aus drei Beobachtungen auf die Lösung zweier verschiedener Aufgaben:

Erstens: *auf irgend eine Weise eine genäherte Bahn zu finden.*

Zweitens: *diese Bahn so zu „verbessern", daß sie den Beobachtungen so gut als möglich Genüge leistet.*

Ist die Bahn schon auf irgend eine Weise genähert bekannt, so entfällt die erste Aufgabe.

Im besonderen gelangt man zur Lösung der ersten Aufgabe, indem man anstatt der beiden Angaben, die der mittlere beobachtete geozentrische Ort liefert, zwei andere benutzt (vgl. Th. m. art. 119); als solche stellt man sich am einfachsten die Abstände des Planeten von der Erde im ersten und dritten Ort vor. Das wichtigste ist also zunächst die Auffindung genäherter Werte dieser beiden Größen und dies ist auf zwei wesentlich verschiedenen Wegen möglich, entweder durch eine reine Versuchsmethode, die im ersten Grunde auf Erraten beruht, oder durch systematisches Vorgehen.

Die reine Versuchsmethode besteht in irgend einer willkürlichen Annahme über die beiden Stücke, auf Grund deren man die Bahn berechnet und die Abweichung von den Beobachtungen feststellt; durch weitere Versuche kann man zu brauchbaren Näherungswerten der beiden Stücke und damit der Bahnelemente gelangen. Da aber die anzunehmende Hypothese sich auf zwei Stücke erstreckt, so werden die Versuche sich in der Regel als praktisch undurchführbar erweisen; daher bemüht man sich, eine genäherte Beziehung zwischen beiden Stücken aufzufinden, die sich durch bekannte Größen ausdrückt, so daß man nur eines willkürlich zu wählen braucht.

So leitet Olbers in seiner Methode der Kometenbahnbestimmung die genäherte Gleichung[1])

$$1)\qquad \delta'' = M\delta$$

ab, wo

$$2)\qquad M = \frac{\operatorname{tg}\beta' \sin(\alpha - L') - \operatorname{tg}\beta \sin(\alpha' - L')}{\operatorname{tg}\beta'' \sin(\alpha' - L') - \operatorname{tg}\beta' \sin(\alpha'' - L')} \cdot \frac{t'' - t'}{t' - t}$$

1) Es bedeutet: δ, δ'' die kurtierten Abstände von der Erde im ersten und dritten Ort,
α, α', α'' die drei beobachteten geozentrischen Längen,
β, β', β'' „ „ „ „ Breiten,
t, t', t'' die drei Beobachtungszeiten,
L' die heliozentrische Länge der Erde im mittleren Ort.

aus den Beobachtungen bekannt ist. Diese Formel beruht auf der genähert giltigen Voraussetzung, daß die Sehnen zwischen den beiden Örtern des beobachteten Körpers und ebenso die zwischen den beiden Örtern der Erde von den Radienvektoren der mittleren Beobachtung im Verhältnis der Zwischenzeiten geschnitten werden, und Olbers leitet sie aus dieser Voraussetzung ab [1]).

3. Übersicht über Gauß' Methoden zur genäherten Bestimmung der Abstände von der Erde (δ und δ'') im ersten und dritten Ort.

Die Olberssche Gleichung finden wir auch in der mehrfach erwähnten Notiz des Handbuchs Bb; sie ist dort von Gauß mit II bezeichnet.

Ebenso finden wir sie in der Summarischen Übersicht im 5. Artikel (Werke VI, S. 157), wo sie aber nur beiläufig abgeleitet und nicht benutzt wird. Auch im Handbuch Bb benutzt Gauß diese Gleichung nicht wie Olbers, der einen willkürlichen Wert von δ annimmt, daraus δ'' berechnet und dann auf Grund dieser beiden hypothetischen Werte die Bahn bestimmt. Gauß geht vielmehr bereits bei seinen ersten Bahnbestimmungen auf dem zweiten erwähnten systematischen Wege vor, indem er direkt aus den Beobachtungen genäherte Werte für δ und δ'' zu finden versucht, auf die er dann die Bahnbestimmung aufbaut. Und zwar findet er genäherte Werte für diese beiden Größen, indem er zunächst einen solchen für den kurtierten Abstand des Planeten δ' in der mittleren Beobachtung sucht.

Zu diesem Zwecke benutzt er eben seine genähert giltige Hauptgleichung [2])

$$3)\qquad \frac{R'}{\delta'}\left(1-\frac{R'^3}{r'^3}\right) = \frac{\operatorname{tg}\beta\sin(\alpha''-\alpha')-\operatorname{tg}\beta'\sin(\alpha''-\alpha)+\operatorname{tg}\beta''\sin(\alpha'-\alpha)}{\operatorname{tg}\beta\sin(L'-\alpha'')-\operatorname{tg}\beta''\sin(L'-\alpha)}\cdot\frac{2}{(M'-M)(M''-M')}.$$

1) Wilhelm Olbers, Sein Leben und seine Werke, herausgegeben von C. Schilling, I. Band, Berlin 1894, S. 30.

2) Es bedeutet: R' den Abstand der Erde von der Sonne in der mittleren Beobachtung,
r' den Abstand des Planeten von der Sonne in der mittleren Beobachtung,
M, M', M'' die mittleren Längen der Erde in den drei Beobachtungen.

Über die Ableitung der Gleichung siehe S. 18 ff.

Im Handbuch Bb stellt Gauß neben diese die Beziehung zwischen δ' und r', die sich unmittelbar aus dem Dreieck Sonne-Erde-Planet ergibt, nämlich

$$4)\qquad \frac{r'}{\delta'} = \sqrt{1 + \mathrm{tg}^2\beta' + \frac{R'^2}{\delta'^2} - 2\frac{R'}{\delta'}\cos(L' - \alpha')}.$$

Aus beiden Gleichungen bestimmt er durch Versuche, die sehr schnell zum Ziele führen, δ', während r' nicht weiter in betracht kommt. Um hieraus genäherte Werte für δ und δ'' zu finden, geht er im Handbuch Bb von der Annahme aus, daß die Logarithmen der drei Abstände von der Erde sich proportional den Zwischenzeiten ändern, also von der Beziehung

$$\frac{\log\delta'' - \log\delta'}{t'' - t'} = \frac{\log\delta' - \log\delta}{t' - t},$$

aus der folgt

$$5)\qquad \begin{aligned} \log\delta &= \log\delta' - \frac{t' - t}{t'' - t}\log\frac{\delta''}{\delta} \\ \log\delta'' &= \log\delta' + \frac{t'' - t'}{t'' - t}\log\frac{\delta''}{\delta}. \end{aligned}$$

Da der Quotient $\frac{\delta''}{\delta}$ aus der Olbersschen Gleichung:

$$6)\qquad \frac{\delta''}{\delta} = \frac{\mathrm{tg}\,\beta'\sin(L' - \alpha) - \mathrm{tg}\,\beta\,\sin(L' - \alpha')}{\mathrm{tg}\,\beta''\sin(L' - \alpha') - \mathrm{tg}\,\beta'\sin(L' - \alpha'')}\cdot\frac{t'' - t'}{t' - t}$$

bekannt ist, so lassen sich die Werte von δ und δ'' berechnen.

Dies ist das älteste Gaußsche Verfahren, wie es uns in der Nachlaßnotiz aus dem Handbuch Bb entgegentritt. In der Summarischen Übersicht hat er diese Methode bereits verfeinert, indem er anstelle der vorigen Gleichung die beiden Gleichungen (Werke VI, S. 156, Gleichungen 5 und 6)

$$7)\qquad \begin{aligned} \delta &= -\frac{f'}{f}\cdot\frac{\mathrm{tg}\,\beta'\sin(L' - \alpha'') - \mathrm{tg}\,\beta''\sin(L' - \alpha')}{\mathrm{tg}\,\beta\,\sin(L' - \alpha'') - \mathrm{tg}\,\beta''\sin(L' - \alpha)}\,\delta' \\ \delta'' &= -\frac{f'}{f''}\cdot\frac{\mathrm{tg}\,\beta\sin(L' - \alpha') - \mathrm{tg}\,\beta'\sin(L' - \alpha)}{\mathrm{tg}\,\beta\sin(L' - \alpha'') - \mathrm{tg}\,\beta''\sin(L' - \alpha)}\,\delta' \end{aligned}$$

anwendet [1]).

Er schreibt hier aber $\frac{g}{f}\cdot\frac{f'}{g'}\cdot\frac{t'' - t}{t'' - t'}$ für $-\frac{f'}{f}$ und $\frac{g''}{f''}\cdot\frac{f'}{g'}\cdot\frac{t'' - t}{t' - t}$ für $-\frac{f'}{f''}$ und nimmt $\frac{f}{g}, \frac{f'}{g'}, \frac{f''}{g''}$ gleich Eins, solange er keinen

1) Es bedeutet: f, f', f'' die Dreiecksflächen zwischen den Radienvektoren der zweiten und dritten, der ersten und dritten, der ersten und zweiten Beobachtung,
g, g', g'' die entsprechenden Sektoren.

besseren Näherungswert für diese Verhältnisse hat. Auf eine Anfrage von Olbers, warum er hier nicht gleich

$$f : f' : f'' = (t''-t') : (t''-t) : (t'-t)$$

setzt, antwortet er im Briefe an Olbers vom 21. Sept. 1802:

„Den Koeffizienten $-\frac{f'}{f}$ würde ich unter die Form

$$-\frac{f'}{f+f''} \cdot \frac{f+f''}{f} = \frac{1}{1-\frac{f+f'+f''}{f'}} \cdot \frac{f+f''}{f}$$

setzen. Der erste Faktor läßt sich sehr nahe bestimmen [[1])] und den zweiten kann man für den Fall, wo die Zwischenzeiten gleich sind, ohne Bedenken $= \frac{t''-t}{t''-t'}$ setzen. Mir scheint, daß man so in diesem Fall der Wahrheit um eine Dimension näher komme, als wenn man gleich $-\frac{f'}{f} = \frac{t''-t}{t''-t'}$ setzte. Ich gestehe indeß, daß ich bisher mich jener Korrektion noch nicht bedient, und mich auch ohne dieselbe ebenso gut befunden habe“.

Der Quotient der beiden Gleichungen ergibt übrigens die Olberssche Gleichung, wie Gauß auch in der Summarischen Übersicht erwähnt.

In der Theoria motus finden wir, wie oben bereits gesagt, die Fundamentalgleichung 3) nicht mehr in der gleichen Form vor. Gauß stellt hier nicht mehr die Anfangswerte der kurtierten Abstände δ und δ'' an die Spitze, sondern er führt als solche die von ihm mit P und Q bezeichneten Größen ein, für die er mit einer Hypothese über das Verhältnis vom Sektor zum Dreieck, die mit der bei Aufstellung der Gleichungen 7) gebrauchten inhaltsgleich ist, die Näherungswerte

$$P = \frac{\Theta''}{\Theta}, \quad Q = \Theta\Theta''$$

ansetzt, deren Erklärung im folgenden (S. 21) gegeben wird. Aus P und Q findet Gauß dann die beiden Größen δ und δ'', sodaß man allerdings die Sache auch so auffassen kann, als ob er diese letzteren Werte seiner Bahnbestimmung zugrunde legt.

---

1) Es ist nämlich näherungsweise $\frac{f+f'+f''}{f'} = -\frac{1}{2r'^3}(M'-M)(M''-M')$.

4. Ableitung von Gauß' Hauptgleichung zur genäherten Bestimmung von δ' und Zusammenhang mit der Theoria motus.

Es treten uns hier also im ganzen drei Methoden der Bahnbestimmung, insbesondere der Bestimmung von δ und δ'' entgegen, die des Handbuchs Bb, die der Summarischen Übersicht und die vervollkommnete der Theoria motus.

Um den Zusammenhang der älteren Methoden mit denen der Theoria motus zu zeigen, mögen die Fundamentalgleichungen zur Bestimmung des Abstandes δ' von der Erde, auf denen einerseits die beiden älteren Methoden unmittelbar, andererseits die der Theoria motus mittelbar beruht, neben einander entwickelt werden. Wir benutzen dabei hier die Bezeichnungen der Theoria motus und schließen uns auch in der Entwicklung an die letztere an; die Ableitung, die Gauß in der Summarischen Übersicht gibt, erscheint verwickelter.

In der Theoria motus, art. 112, leitet Gauß die bekannten Gleichungen[1]):

8)
$$\begin{aligned} 0 &= nx - n'x' + n''x'' \\ 0 &= ny - n'y' + n''y'' \\ 0 &= nz - n'z' + n''z'' \end{aligned}$$

ab; ihre Ableitung finden wir auch in der Summarischen Übersicht, art. 4, jedoch an dieser Stelle auch die entsprechenden Gleichungen für die Erde:

9)
$$\begin{aligned} 0 &= NX - N'X' + N''X'' \\ 0 &= NY - N'Y' + N''Y'' \\ 0 &= NZ - N'Z' + N''Z'', \end{aligned}$$

die in der Theoria motus nicht angewendet werden; hierin liegt der wesentliche Unterschied beider Methoden.

---

1) Es bedeutet: x, y, z usw. die heliozentrischen Koordinaten des Planeten in den drei Örtern,
n, n', n'' die doppelten Dreiecksflächen zwischen den Radienvektoren des zweiten und dritten, des ersten und dritten, des ersten und zweiten Ortes,
X, Y, Z usw., N, N', N'' die entsprechenden Größen für die Erde,
D, D', D'' die kurtierten Abstände der Erde von der Sonne,
L, L', L'' die heliozentrischen Längen der Erde,
B, B', B'' „ „ Breiten „ „

Setzt man in den Gleichungen 8) die Werte

$$10)\qquad \begin{aligned} x &= \delta\cos\alpha + D\cos L \\ y &= \delta\sin\alpha + D\sin L \\ z &= \delta\,\mathrm{tg}\,\beta + D\,\mathrm{tg}\,B \end{aligned}$$

und entsprechend die von $x', y', z', x'', y'', z''$ ein, so folgen die Gleichungen (1)—(3) des art. 112 der Theoria motus:

$$11)\qquad \begin{aligned} 0 &= n(\delta\cos\alpha + D\cos L) - n'(\delta'\cos\alpha' + D'\cos L') + n''(\delta''\cos\alpha'' + D''\cos L'') \\ 0 &= n(\delta\sin\alpha + D\sin L) - n'(\delta'\sin\alpha' + D'\sin L') + n''(\delta''\sin\alpha'' + D''\sin L'') \\ 0 &= n(\delta\,\mathrm{tg}\,\beta + D\,\mathrm{tg}\,B) - n'(\delta'\,\mathrm{tg}\,\beta' + D'\,\mathrm{tg}\,B') + n''(\delta''\,\mathrm{tg}\,\beta'' + D''\,\mathrm{tg}\,B''). \end{aligned}$$

Eliminiert man aus diesen δ und δ'', indem man der Reihe nach mit

$$-\mathrm{tg}\,\beta\sin\alpha'' + \mathrm{tg}\,\beta''\sin\alpha,\ \ \mathrm{tg}\,\beta\cos\alpha'' - \mathrm{tg}\,\beta''\cos\alpha,\ \ \sin(\alpha''-\alpha)$$

multipliziert und addiert, so wird:

$$12)\qquad a\delta' = b + \frac{n}{n'}c + \frac{n''}{n'}d,$$

wo

$$13)\qquad \begin{aligned} a &= \mathrm{tg}\,\beta\sin(\alpha''-\alpha') + \mathrm{tg}\,\beta'\sin(\alpha-\alpha'') + \mathrm{tg}\,\beta''\sin(\alpha'-\alpha) \\ b &= \ \ D'\left\{\mathrm{tg}\,\beta\sin(L'-\alpha'') + \mathrm{tg}\,B'\sin(\alpha''-\alpha) - \mathrm{tg}\,\beta''\sin(L'-\alpha)\right\} \\ c &= -D\left\{\mathrm{tg}\,\beta\sin(L-\alpha'') - \mathrm{tg}\,\beta''\sin(L-\alpha) + \mathrm{tg}\,B\sin(\alpha''-\alpha)\right\} \\ d &= -D''\left\{\mathrm{tg}\,\beta\sin(L''-\alpha'') - \mathrm{tg}\,\beta''\sin(L''-\alpha) + \mathrm{tg}\,B''\sin(\alpha''-\alpha)\right\} \end{aligned}$$

bekannt sind.

Diese in art. 132 der Theoria motus abgeleitete Gleichung 12) ist die strenge Bestimmungsgleichung für δ', in der die noch unbekannten Größen n, n', n'' zur ersten Bestimmung durch genäherte Werte ersetzt werden. Gauß zeigt (Theoria motus, art. 132—133), daß man bis auf Größen 4. Ordnung richtig schreiben kann [1])

$$14)\qquad a\delta' = b + \frac{c\Theta + d\Theta''}{\Theta'}\left(1 + \frac{\Theta\Theta''}{2r'^3}\right),$$

weil die Beziehungen

$$\frac{cn + dn''}{n + n''} = \frac{c\Theta + d\Theta''}{\Theta'},\qquad \frac{n + n''}{n'} = 1 + \frac{\Theta\Theta''}{2r'^3}$$

bis auf Größen 4. Ordnung richtig sind. Aus der Gleichung 14)

1) Es bedeuten $\Theta, \Theta', \Theta''$ die Zwischenzeiten zwischen der zweiten und dritten, der ersten und dritten, der ersten und zweiten Beobachtung, multipliziert mit der Gaußschen Konstante k.

läßt sich nun durch Heranziehung der ähnlichen Gleichung für die Erdbewegung die Fundamentalgleichung der Summarischen Übersicht und der Nachlaßnotiz aus Handbuch Bb ableiten. Die Gleichungen 9) werden nämlich

15)
$$\begin{aligned} ND\cos L - N'D'\cos L' + N''D''\cos L'' &= 0 \\ ND\sin L - N'D'\sin L' + N''D''\sin L'' &= 0 \\ ND\,\mathrm{tg}\,B - N'D'\,\mathrm{tg}\,B' + N''D''\,\mathrm{tg}\,B'' &= 0. \end{aligned}$$

Multipliziert man diese mit denselben Größen wie oben die Gleichungen 11) und addiert, so erhält man ähnlich der Gleichung 12):

16)
$$b = -\frac{N}{N'}c - \frac{N''}{N'}d.$$

Auch hier kann man genähert setzen

$$b = -\frac{c\Theta + d\Theta''}{\Theta'}\left(1 + \frac{\Theta\Theta''}{2R'^3}\right)$$

oder mit der gleichen Genauigkeit

$$\frac{c\Theta + d\Theta''}{\Theta'} = -b\left(1 - \frac{\Theta\Theta''}{2R'^3}\right)$$

und dieser Wert in 14) eingesetzt, gibt wieder bis zur 4. Ordnung genau

17)
$$a\delta' = -b\frac{\Theta\Theta''}{2}\left(\frac{1}{r'^3} - \frac{1}{R'^3}\right)$$

oder

$$\frac{R'}{\delta'}\left(1 - \frac{R'^3}{r'^3}\right) = \frac{a}{b}\cdot\frac{2R'^4}{\Theta\Theta''};$$

diese Gleichung endlich geht in die Fundamentalgleichung 3) über, wenn man die Breite der Erde vernachlässigt, also

$$B' = 0, \quad R' = D'$$

setzt und bedenkt, daß

$$\begin{aligned} M' - M &= \frac{k}{R'^{\frac{3}{2}}}(t' - t) = \frac{\Theta''}{R'^{\frac{3}{2}}} \\ M'' - M' &= \frac{k}{R'^{\frac{3}{2}}}(t'' - t') = \frac{\Theta}{R'^{\frac{3}{2}}} \end{aligned}$$

genommen werden kann.

5. Methode der Theoria motus zur genäherten Bestimmung des Abstandes δ' im mittleren Ort.

Wie schon erwähnt, stellt Gauß in seinen älteren Methoden neben die Hauptgleichung 3) die Beziehung 4) und bestimmt δ'

aus beiden durch Proben oder wie er es nennt, durch die indirekte Methode. Er sagt in art. 6 der Summarischen Übersicht „Die indirekte Methode ist hier bei weitem die bequemste; man kommt nach wenigen Versuchen, wofür sich leicht zweckmäßige Vorschriften geben lassen, sehr schnell zum Ziele".

In der Theoria motus hat er aber diese indirekte Methode verlassen. Dort handeln die art. 139—141 von der Bestimmung von δ', oder vielmehr von der von r', da hier die Abstände von der Sonne anstelle derer von der Erde benutzt werden, und zwar finden wir hier eine geometrische Lösung der Aufgabe, bei der der Zusammenhang mit der ursprünglichen Hauptgleichung 3) und selbst der mit der in der Theoria motus auftretenden entsprechenden Gleichung 14) nicht mehr zu erkennen ist. Darum haben bereits Encke[1]) und Klinkerfues[2]) eine von der Gaußschen abweichende Ableitung gegeben.

Wir wollen, um den Zusammenhang zu zeigen, die Gleichungen der Theoria motus art. 141 möglichst kurz aus der Hauptgleichung und zwar in ihrer strengen Form 12) ableiten. Diese erstere lautet mit den Gaußschen Bezeichnungen P und Q (Theoria motus, art. 134)

18) $$a\delta' = b + \frac{c+dP}{1+P}\left(1+\frac{Q}{2r'^3}\right),$$

wo

19) $$P = \frac{n''}{n}, \quad Q = 2\left(\frac{n+n''}{n'}-1\right)r'^3.$$

Nennt man im Dreieck Sonne-Erde-Planet, dessen drei Seiten R', r', Δ' sind, den Winkel am Planeten z und den aus der Beobachtung bekannten Winkel an der Erde $180^\circ - \lambda$, so ist

$$\Delta' = \delta' \cos\beta' = \frac{R'\sin(\lambda - z)}{\sin z}, \quad r' = \frac{R'\sin\lambda}{\sin z}.$$

Es wird also die Hauptgleichung

$$\frac{aR'}{\cos\beta'}\sin(\lambda - z) = \left(b+\frac{c+dP}{1+P}\right)\sin z + \frac{c+dP}{1+P}\,\frac{Q\sin^4 z}{2R'^3\sin^3\lambda}$$

oder

$$\frac{aR'}{\cos\beta'}\sin\lambda\cos z - \left(b+\frac{aR'}{\cos\beta'}\cos\lambda\right)\sin z$$
$$= \frac{c+dP}{1+P}\left(\sin z + \frac{Q\sin^4 z}{2R'^3\sin^3\lambda}\right).$$

1) Über die Bestimmung einer elliptischen Bahn aus drei vollständigen Beobachtungen, art. 9—10. Berliner Astronomisches Jahrbuch für 1854.

2) Klinkerfues, Theoretische Astronomie, 48. Vorlesung, Braunschweig 1871, zweite Auflage, herausgegeben von H. Buchholz, Braunschweig 1899.

Setzt man

$$-\frac{aR'}{\cos\beta'}\sin\lambda = g\sin\sigma, \quad -\left(b+\frac{aR'}{\cos\beta'}\cos\lambda\right) = g\cos\sigma,$$

so wird

$$g\cdot\frac{1+P}{c+dP}\sin(z-\sigma) = \sin z+\frac{Q\sin^4 z}{2R'^3\sin^3\lambda},$$

welche Gleichung mit der des art. 141 der Theoria motus

$$\sin z+\frac{Q\sin^4 z}{2R'^3\sin^3\lambda} = b\frac{P+1}{P+a}\sin(z-\sigma)$$

identisch ist, wo aber die mit a, b bezeichneten Größen nicht gleich den oben und im art. 132 der Theoria motus ebenso bezeichneten sind. Gauß hat die vorstehende Gleichung im Jahre 1806 aufgestellt; sie findet sich nämlich im Handbuch Bd, S. 92. Durch Einführung des Winkels σ verschwinden die beiden Größen δ' und r' aus der Hauptgleichung und die Arbeit reduziert sich auf die Bestimmung dieses Winkels, zu welchem Zweck Gauß im art. 141 der Gleichung noch die bequemere Form

20) $$cQ\sin\omega\sin^4 z = \sin(z-\omega-\sigma)$$

gibt, wo

$$c = \frac{1}{2R'^3\sin^3\lambda\sin\sigma} \quad \text{und} \quad \operatorname{tg}\omega = \frac{\sin\sigma}{b\dfrac{P+1}{P+a}-\cos\sigma}$$

gesetzt ist. In der Gleichung 20) ist z allein unbekannt, wenn man von den im Verlauf der Annäherungen zu bestimmenden Größen P und Q absieht, von denen die erstere im Winkel ω vorkommt und für die Gauß in der ersten Annäherung setzt:

$$P = \frac{\Theta''}{\Theta}, \quad Q = \Theta\Theta''.$$

Es ergibt sich sodann r' aus der Gleichung

21) $$r' = \frac{R'\sin\lambda}{\sin z}.$$

6. Bestimmung von δ und δ'' aus δ' bezw. von r und r'' aus r'.

Nachdem r' bezw. δ' gefunden sind, hat man jetzt r und r'' bezw. δ und δ'' zu berechnen. In der ältesten Nachlaßnotiz bedient sich Gauß der Beziehungen 5) und der Olbersschen Gleichung 6), um δ und δ'' aus δ' zu finden, in der Summarischen Übersicht der schon

verfeinerten Gleichungen 7). In der Theoria motus, wo die r statt der δ angewandt sind, erscheint das weiter verbesserte Verfahren verwickelter: Gauß berechnet hier zunächst die Ausdrücke (art. 143—144)

$$22)\qquad \frac{n'r'}{n} = \frac{(P+a)R'\sin\delta'}{b\sin(z-\sigma)}, \quad \frac{n'r'}{n''} = \frac{1}{P}\frac{n'r'}{n},$$

$$23)\qquad \begin{aligned} r\sin\zeta &= \frac{n'r'}{n}\frac{\sin\varepsilon}{\sin\varepsilon'}\sin(z+A'D-\delta') = p\\ r\cos\zeta &= \varkappa(\lambda p-1) = q\\ r''\sin\zeta'' &= \frac{n'r'}{n''}\frac{\sin\varepsilon''}{\sin\varepsilon'}\sin(z+A'D''-\delta') = p''\\ r''\cos\zeta'' &= \varkappa''(\lambda''p''-1) = q'', \end{aligned}$$

wobei die Größen ε, ε', ε'', $A'D$, $A'D''$, δ' aus der vorhergehenden Rechnung (art. 136—137) bekannt sind und auch $\varkappa$, $\varkappa''$, λ, λ'' sich durch bekannte Größen ausdrücken.

Er findet so neben r und r'' noch die beiden Winkel ζ und ζ'', die mit Hilfe der Beziehungen [1])

$$24)\qquad \begin{aligned} \sin f'\sin\frac{u''+u}{2} &= \sin\frac{\varepsilon'}{2}\sin\frac{\zeta+\zeta''}{2}\\ \sin f'\cos\frac{u''+u}{2} &= \cos\frac{\varepsilon'}{2}\sin\frac{\zeta-\zeta''}{2}\\ \cos f'\sin\frac{u''-u}{2} &= \sin\frac{\varepsilon'}{2}\cos\frac{\zeta+\zeta''}{2}\\ \cos f'\cos\frac{u''-u}{2} &= \cos\frac{\varepsilon'}{2}\cos\frac{\zeta-\zeta''}{2} \end{aligned}$$

auch gleich f' und die Winkel u, u'' liefern; f' dient zur Verbesserung der Werte von P und Q, während u und u'' erst später bei genauerer Bestimmung der Bahn gebraucht werden. Vorarbeiten hierzu finden sich im Handbuch Bd, S. 92.

7. Bestimmung einer genäherten Bahn aus hypothetischen Werten von δ und δ'' oder von i und Ω, nach den älteren Methoden.

Sind einmal außer den beobachteten Örtern zwei Stücke, und zwar insbesondere zwei Abstände des Planeten von der Erde

1) Es bedeutet hier $f' = \frac{v''-v}{2}$ die halbe Differenz der Anomalien (oder Längen in der Bahn) im ersten und dritten Ort.

Die Winkel ζ, ζ'', u, u'' sind Hilfswinkel, deren geometrische Bedeutung aus der Figur 4 und dem Artikel 149 der Th. m. hervorgeht.

(oder der Sonne) genähert bekannt, so kann man daraus eine genäherte Bahn berechnen und die Lösung der zweiten Aufgabe, nämlich diese Bahn zu „verbessern", entweder daran anschließen oder auch unmittelbar damit vereinigen. Auch hier kann man zwei prinzipiell verschiedene Wege einschlagen, indem man entweder eine reine Versuchsmethode gebraucht oder systematisch vorgeht.

Den ersteren Weg benutzt Gauß stets bei seinen ersten Bahnbestimmungen, und zwar so, daß er eine genäherte Bahn nicht nur aus den gefundenen hypothetischen Werten der Abstände, sondern auch aus zwei weiteren benachbarten Wertepaaren berechnet und sodann diejenigen Werte interpoliert, welche die Beobachtungen darstellen. Er geht also von drei angenommenen Wertepaaren, etwa: 1. δ, δ''; 2. $\delta + \varepsilon, \delta''$; 3. $\delta, \delta'' + \varepsilon''$ aus und berechnet aus allen dreien je eine Bahn. Die Vergleichung dieser Bahnen mit den Beobachtungen gibt die Mittel an die Hand, um auf die richtigen Bahnelemente zu schließen. (Vgl. Theoria motus, art. 120—122.) Die Einzelheiten des Verfahrens können dabei sehr verschieden gewählt werden und an verschiedenen Stellen gibt Gauß hierfür verschiedene Methoden an.

Im Handbuch Bb, sowie in der Summarischen Übersicht, art. 9, unter I., erwähnt Gauß zunächst seine „erste Verbesserungsmethode", die eben darin besteht, daß man aus den drei Hypothesen für δ und δ'' drei Bahnen berechnet, die die beiden äußeren Beobachtungsörter scharf darstellen; sodann berechnet man aus den drei Elementensystemen die mittlere Beobachtung und schließt aus den Abweichungen dieser durch Interpolation auf diejenigen Werte der Abstände, oder auf diejenigen Werte der Elemente, welche auch diesen mittleren Ort so scharf wie möglich darstellen; im Bedarfsfalle wiederholt man das gleiche Verfahren, bis die erwünschte Genauigkeit erreicht ist.

Zur Bestimmung der Bahn aus den verschiedenen Hypothesen dienen die „Vorschriften zur Berechnung der Elemente, aus zweien geozentrischen Örtern, der Zwischenzeit, und den zugehörigen Abständen" in der Nachlaßnotiz aus dem Handbuch Bb, vgl. Werke XI 1. Wie in dieser ganzen Nachlaßnotiz, so gibt Gauß hier eine Fülle von Formeln zur Lösung der vorgesetzten Aufgabe, aus denen der Rechner sich die heraussuchen kann, die ihm am bequemsten scheinen. Wie oben schon gesagt, dürfte hiermit die Tagebuchnotiz Nr. 121 vom Oktober 1801 in Verbindung stehen. Leider ist aus jener Zeit im Nachlaß nichts erhalten, was uns Aufschluß über die Entstehung dieser Formeln geben könnte; auch unter den vielen in den Schedae Ag und Ah vorhandenen numerischen Rechnungen findet sich keine,

die von hypothetischen Werten der Abstände δ und δ'' ausgehen. Nur in der späteren (1802) Scheda Ai ist auf S. 14 und anscheinend S. 20 f. ein Bruchstück einer solchen Rechnung für Ceres erhalten. Ebenso beruhen mehrere der Bahnbestimmungen für Pallas (von denen aber die allerersten ebenfalls nicht im Nachlaß erhalten sind) auf dieser Methode (Scheda Ak zu Anfang und besonders Al). Dagegen gehen die ältesten erhaltenen Rechnungen, die sich auf Ceres beziehen, und einige auf Pallas bezügliche, von hypothetischen Werten für die Neigung i und die Knotenlänge Ω aus. Nach den Formeln der Notiz im Handbuch Bb (vgl. Werke XI 1) findet man nun freilich, wenn δ und δ'' bekannt sind, nach dem ersten Teil der „Vorschriften, zur Berechnung der Elemente aus zweien geozentrischen Örtern, der Zwischenzeit, und den zugehörigen Abständen", zu denen Gauß kein numerisches Beispiel gibt, auch i und Ω; aber nicht nur diese allein, sondern auch die heliozentrischen Längen und Breiten, die wahren Längen in der Bahn und die Abstände von der Sonne. In seinen numerischen Rechnungen löst Gauß aber meist die Aufgabe so, daß er hypothetische Werte von i und Ω an die Spitze stellt, daraus als erstes Geschäft die Radienvektoren und die Längen in der Bahn für alle drei Beobachtungen berechnet, sodann die Bahn aus den beiden äußeren Örtern in der Bahn bestimmt und schließlich aus den gefundenen Bahnelementen den mittleren Ort (ebenfalls in der Bahn) berechnet und seine Abweichung von der Beobachtung feststellt. Indem er dies auch hier für drei Hypothesen, etwa: 1. i, Ω; 2. $i+\varepsilon_1, \Omega$; 3. $i, \Omega+\varepsilon_2$ ausführt, erhält er mit den genannten Abweichungen die nötigen Daten, um verbesserte Werte von i und Ω zu interpolieren, mit denen er dieselbe Rechnung wiederholt. Die hierauf bezüglichen Nachlaßnotizen werden in Band XI 1 der Werke veröffentlicht werden.

Die von hypothetischen Werten von i und Ω ausgehende Methode nennt Gauß sowohl in der Nachlaßnotiz als in der Summarischen Übersicht seine „Zweite Verbesserungsmethode". Bei ihr ist die Aufgabe zu lösen: *Aus zwei geozentrischen Örtern, der Neigung und der Länge des Knotens die Bahn zu berechnen*, und zwar zerfällt diese Aufgabe in zwei Einzelaufgaben, Erstens: *Aus einem geozentrischen Orte, der Neigung und der Länge des Knotens, die heliozentrische Länge in der Bahn und den Abstand von der Sonne zu finden*, welche Aufgabe für jede der beiden äußern Beobachtungen zu lösen ist, und Zweitens: *Aus zwei heliozentrischen Längen in der Bahn und den zugehörigen Abständen von der Sonne die Bahnelemente zu finden.*

Die Lösung der ersten Einzelaufgabe finden wir in der ein-

fachsten Form im Handbuch Bb. Gauß hat sich aber mit ihr anscheinend vielfach beschäftigt. In der Scheda Ag, S. 5, ist sie mit Beispielen in derselben Weise gelöst wie an der eben erwähnten Stelle; weiter finden wir aber in der Scheda Ag, S. 56, und im Handbuch Bb, S. 15, eine weitere Ausführung derselben Aufgabe zum Zwecke, ihre Auflösung für logarithmisches Rechnen geschmeidiger zu gestalten, und endlich hat er diese letztere Lösung in der Monatl. Corr., Juni 1802 (Werke VI, S. 87) veröffentlicht.

Die zweite Einzelaufgabe bietet reizvollere Einzelheiten. Die Lösung kann auch nicht explizit erfolgen. Wenn die Bahnelemente noch völlig unbekannt sind, so geht Gauß von einem Näherungswert des Parameters der Bahn p aus und der Kunstgriff, durch den er sich diesen Näherungswert verschafft, verdient besonders hervorgehoben zu werden. Er findet sich angedeutet im Handbuch Bb, ausführlicher dargestellt in der Summarischen Übersicht (Werke VI, S. 161) und besteht im Grunde aus einer mechanischen Quadratur:

Der Sektor g' zwischen den beiden äußern Beobachtungen ist, wenn man die Masse des Planeten vernachlässigt,

$$g' = \frac{k}{2}\sqrt{p}\,.(t''-t),$$

wo k die Gaußsche Konstante bedeutet. Nimmt man den Erdbahnhalbmesser als Einheit, also die mittlere Bewegung der Erde gleich k, so ist

$$g' = \frac{\sqrt{p}}{2}(U''-U),$$

wenn U, U'' die mittleren Längen der Sonne sind; also wird

$$\sqrt{p} = \frac{2g'}{U''-U}.$$

Es ist aber, wenn v und v'' die wahren Längen im ersten und dritten Ort bedeuten,

$$2g' = \int_v^{v''} r^2\,dw = \int_v^{v''} \varphi(w)\,dw,$$

wofür man nach der Formel von Cotes genähert setzen kann

$$2g' = \frac{1}{2}(\varphi(v)+\varphi(v''))(v''-v)$$

oder genauer

$$2g' = \left(\frac{1}{6}\varphi(v)+\frac{2}{3}\varphi\left(\frac{v+v''}{2}\right)+\frac{1}{6}\varphi(v'')\right)(v''-v).$$

Es wird also für die erste rohe Annäherung

$$2g' = \frac{1}{2}(r^2 + r''^2)(v'' - v)$$

und damit

25)
$$\sqrt{p} = \frac{r^2 + r''^2}{2} \cdot \frac{v'' - v}{U'' - U}$$

und genauer

$$2g' = \frac{r^2 + r''^2 + 4\mathfrak{R}}{6}(v'' - v),$$

also

26)
$$\sqrt{p} = \frac{r^2 + r''^2 + 4\mathfrak{R}}{6} \cdot \frac{v'' - v}{U'' - U},$$

wenn $\mathfrak{R}$ der zur Länge $\frac{v + v''}{2}$ gehörende Radiusvektor des Planeten ist. Für den letzteren leitet man aus den Gleichungen

$$\frac{p}{r} - 1 = e \cos(v - \pi)$$

$$\frac{p}{r''} - 1 = e \cos(v'' - \pi)$$

$$\frac{p}{\mathfrak{R}} - 1 = e \cos\left(\frac{v + v''}{2} - \pi\right)$$

durch Elimination von e und π den Wert

27)
$$\frac{1}{\mathfrak{R}} = \frac{\frac{1}{2}\left(\frac{1}{r} + \frac{1}{r''}\right)}{\cos\frac{v'' - v}{2}} - \frac{2\sin^2\frac{v'' - v}{4}}{p\cos\frac{v'' - v}{2}}$$

ab. Nachdem man aus der Formel 25) p genähert gefunden, ergibt sich $\mathfrak{R}$ und damit aus 26) auch ein genauerer Wert von p.

Diese Bestimmungsmethode von p wendet Gauß in einer in Band XI 1 der Werke abzudruckenden numerischen Bahnbestimmung der Ceres an. Aus dem vorausgesetzten Werte von p ergibt sich dann leicht die ganze Bahn (vgl. die Formeln in der Nachlaßnotiz aus dem Handbuch Bb und die Bahnbestimmung der Ceres in Bd. XI 1).

In der Theoria motus, art. 86 erwähnt er diese Methode auch, ohne daß sie dort praktische Verwendung findet. Im Briefe an Olbers vom 21. Sept. 1802 schreibt Gauß den Wert für den Parameter in einer noch bequemeren Form, die sich auch im Nachlaß in der Schedae Ag, S. 3, vorfindet.

Anstatt die Bahnbestimmung mit einem vorausgesetzten Wert von p vorzunehmen, kann man auch von genäherten Werten der

Exzentrizität oder der Perihellänge ausgehen, wie Gauß im Handbuch Bb und in der Summarischen Übersicht zeigt (vgl. auch art. 80—81 der Theoria motus); indessen ist dies nur angängig, wenn man diese Größen schon genähert kennt.

Im Handbuch Bb gibt Gauß noch eine „Dritte Verbesserungsmethode" an, die, ebenso wie ein Beispiel dazu, in Band XI 1 der Werke zum Abdruck gelangt.

8. Verbesserung der Bahn nach den älteren Methoden.

Hat man jetzt durch Lösung beider Einzelaufgaben die genäherte Bahn gefunden, die unter der für i und Ω oder δ und δ'' gemachten Hypothese die beiden äußeren Beobachtungen scharf darstellt, so berechnet man, wie bereits erwähnt, aus den Bahnelementen den mittleren Ort und stellt seine Abweichung vom beobachteten fest. Hier benutzt nun Gauß bei seinen im Nachlaß erhaltenen Rechnungen zur Bahnbestimmung den mittleren heliozentrischen Ort, während er im Handbuch Bb den geozentrischen dafür empfiehlt; das erstere ist das bequemere.

Nachdem für die drei Hypothesen von i und Ω die entsprechenden Fehler des mittleren Orts gefunden sind, bestimmt Gauß die wahren Werte dieser Größen durch Interpolation; die numerische Ausführung erhellt aus einem im Band XI 1 der Werke zum Abdruck kommenden Beispiel. Weitere Aufzeichnungen darüber finden sich nicht im Nachlaß; in der Theoria motus handeln die art. 120—122 davon.

Im art. 8 der Summarischen Übersicht führt Gauß noch eine andere Verbesserungsmethode an, auf die er „erst bei Veranlassung der Pallas verfiel" und die er dort „mit dem glücklichsten Erfolg" anwandte (Werke VI, S. 163). Sie besteht darin, daß man eine gefundene Bahn, die die äußeren Beobachtungen genau und die mittlere mit gewissen Fehlern darstellt, dadurch verbessert, daß man die ganze Rechnung von der Bestimmung von δ und δ'' an von neuem macht und dazu die äußern Beobachtungen nach ihrem wahren Wert ansetzt, die mittlere jedoch so, daß man an ihr die gefundenen Fehler mit entgegengesetzten Vorzeichen anbringt.

9. Geschichtliches über Gauß' Bahnbestimmungen.

Die älteren Rechnungen zur Bestimmung der Bahn der Ceres, befinden sich, soweit sie erhalten sind, in den Schedae Ag und Ah und stammen aus dem November 1801; die allerersten aus dem Oktober und vielleicht schon aus dem September sind nicht er-

halten. Gauß hat die Elementensysteme, welche er bei weiterer Verbesserung der Bahn fand, numeriert. Die Elementensysteme I—III sind nach den Piazzischen im September-Heft 1801 der Monatl. Corr. veröffentlichten Beobachtungen berechnet; die Berechnung der als I. bezeichneten Elemente (vgl. Werke XI 1) erfolgte nach der „Dritten Verbesserungsmethode", von der er später in der Theoria motus, art. 82 (vgl. auch Artikel 126) sagt, daß sie „orbitae dimensiones eruendi magnam praecisionem nunquam admittet, nisi tria loca heliocentrica intervallis considerabilibus ab invicem distent".

Dabei ist bemerkenswert, daß auch hier von einer Hypothese über i und Ω ausgegangen wird, ohne daß ersichtlich ist, wie die Ausgangswerte ($\Omega = 81^\circ 27' 4'',62$; $\log \operatorname{tg} i = 9.262\,1790$) erlangt worden sind. Gauß hat sie entweder aus den bereits vorliegenden Bahnbestimmungen von Olbers, Burckhardt oder Piazzi entnommen, da selbst die wenig sicheren Kreisbahnbestimmungen gerade für diese beiden Elemente einigermaßen brauchbare Näherungswerte liefern oder er hat sie durch die verlorenen allerersten Versuche der Bahnbestimmung erhalten. Dabei ist aber auffallend, daß Gauß keine runden Anfangswerte für sie annimmt, sondern z. B. in Ω die Sekunden zu $4'',62$ ansetzt. Daß er sie bis auf Hundertel Sekunden überhaupt ansetzt, darf weniger verwundern, weil er ja eine möglichst scharfe Darstellung der Piazzischen Beobachtungen beabsichtigte, wenn auch diese Genauigkeit bei der ersten Bahnbestimmung und mit Rücksicht darauf überflüssig sein mag, daß die Beobachtungen selbst nicht entsprechend genau sind. Aber Gauß liebte es, mit einer übertriebenen Anzahl von Stellen zu rechnen, auch da, wo nicht, wie hier, eine gewisse Begründung dafür gefunden werden kann[1]). Wahrscheinlich aber ist die Ableitung der Elemente I nicht die erste Bahnbestimmung der Ceres, die Gauß ausgeführt hat; jedenfalls sind die Rechnungen, die vor der Scheda Ag liegen, verloren gegangen[2]), wie es ja auch gewiß mit der Entwicklung der Formeln aus dem September und dem Oktober der Fall ist, die wir in den Stücken Bb, Ag und an anderen Stellen des Nachlasses fertig zusammengestellt finden.

Der erhaltene Teil der Rechnung der Elemente I (Werke XI 1) bezieht sich nur auf die Bahnbestimmung aus der ersten Hypothese über i und Ω und schließt mit der Berechnung der Zwischenzeiten zwischen den drei Beobachtungen; es findet sich dann nur

---

1) Vgl. Galle, Gauß als Zahlenrechner, Materialien, Heft IV.

2) Gauß sagt in der Einleitung zur Theoria motus (Werke VII, 1906, S. 8), daß er die erste Bahnbestimmung der Ceres im Oktober 1801 gemacht habe.

noch die Bemerkung: „Durch Interpolation fand man sodann folgende Elemente, welche die beiden äußern Beobachtungen genau, die mittlere mit $+1'',84$ Fehler in der Länge und $-2'',36$ in der Breite [vgl. Werke VI, S. 200] darstellen", worauf die ungeordnet angegebenen Werte der Elemente I folgen, ohne daß angegeben ist, wie diese Interpolation ausgeführt wurde.

Es scheint übrigens, daß die Scheda Ag ursprünglich aus losen Blättern bestand, die erst nach ihrer Benutzung geheftet wurden, so daß das Auftreten von Lücken erklärlich ist. Von der Berechnung der Elemente II ist kaum etwas erhalten.

Nachdem Gauß sie abgeleitet und gesehen hatte, daß beide Elementensysteme I und II die Piazzischen Beobachtungen gut darstellten, sandte er beide an v. Zach, wie aus einem im Gauß-Archiv erhaltenen Brief v. Zachs hervorgeht. Da dieser inzwischen die kleine Abhandlung von Piazzi mit den verbesserten Beobachtungen erhalten hatte, so teilte er die letzteren sofort Gauß mit, der aber schon mit der Berechnung seiner III. Elemente fast fertig war. Diese, sowie die auf Grund der verbesserten Piazzischen Beobachtungen berechneten IV. Elemente, teilte er nun auch v. Zach mit und fügte eine Ephemeride bei, die vom 25. Nov.—31. Dez. reichte. v. Zach veröffentlichte die gesamten bisherigen Resultate von Gauß im Dezember-Heft 1801 der Monatl. Corr. (Werke VI, S. 200). Nachdem v. Zach und viele andere die Ceres auch nach der Gaußschen Ephemeride im Dezember bei meist ungünstigem Wetter vergebens gesucht hatten, konnte der erstere endlich im Februar-Heft 1802 von der glücklichen Wiederauffindung des Planeten berichten. Brieflich hatte er sie Gauß mitgeteilt am 17. Januar 1802. In der Nacht vom 31. Dezember auf den 1. Januar konnte v. Zach nämlich feststellen, daß ein am 7. Dezember von ihm beobachteter verdächtiger Stern die Ceres war. Am 1. Januar fand auch Olbers den Planeten auf; sein Ort stimmte sehr genau mit der Gaußschen Ephemeride überein. Gauß scheint von der Wiederauffindung der Ceres durch Olbers zuerst aus den Zeitungen gehört zu haben und er richtete an diesen am 18. Januar einen Brief, um Olbers' Beobachtungen zu erhalten; mit diesem Brief beginnt der Briefwechsel zwischen beiden, aus dem man, ebenso wie aus den fortgesetzten Mitteilungen v. Zachs über die Beobachtungen und Berechnungen der Ceres in der Monatl. Corr. (Werke VI, von S. 199 an), entnehmen kann, wie die darauf folgende Zeit nun für Gauß fast gänzlich mit neuen Verbesserungen der Ceresbahn ausgefüllt war.

Nach der Wiederauffindung der Ceres, bis zu welcher Gauß

noch ein V. Elementensystem berechnet hatte, machte er sich sogleich daran, die neuen Zachschen Beobachtungen damit zu vergleichen und v. Zach sagt in der Monatl. Corr. vom März 1802: „Als ich dem Dr. Gauß die Nachricht von der glücklichen Auffindung der so sehnlichst erwarteten Ceres, und meine drei ersten Beobachtungen derselben mitgeteilt hatte, so war das erste, was er nach Empfang derselben tat, daß er sie sogleich nach seinen oben angezeigten V. Elementen berechnete. Er fand den Fehler bei der ersten Beobachtung vom 7. Dezember in AR $+24' 8''$; bei der zweiten vom 11. Januar $+30' 53''$; bei der dritten vom 16. Januar $+31' 53''$. Nach seinen IV. (verbesserten) Elementen weichen sie nach einem Überschlage in folgender Ordnung ab: $+14\frac{1}{2}$ Minuten, $+19\frac{3}{4}$ Minuten, $+20\frac{1}{2}$ Minuten. Daß diese IV. Elemente der Wahrheit etwas näher kommen, als die V., hält Dr. Gauß für Zufall; vielleicht ist es aber auch zum Teil Folge der Einwirkung der Planetenstörungen bei den Piazzischen Beobachtungen, besonders auf die Breiten“.

In der Folgezeit ist ein großer Teil von Gauß' Zeit durch Bahnverbesserung und Berechnungen, auch der inzwischen entdeckten Pallas (April 1802), der Juno (Sept. 1804) und der Vesta (März 1807) in Anspruch genommen. Die erste Bahn für Vesta berechnete Gauß in nur 10 Stunden (vgl. Werke VI, S. 285—286). Da er sich außerdem wohl Sorgen um seine Zukunft machte und die Berufungen nach Petersburg (von 1802 an) und nach Landshut (1802) und die Aussicht auf die Berufung nach Göttingen in jene Zeit fielen, und da er sich auch längere Zeit auf der Sternwarte Seeberg (1803) aufhielt, so dauerte nicht nur die Unterbrechung seiner mathematischen Untersuchungen an, sondern auch die weitere Ausfeilung seiner Methoden zur Bahnbestimmung und die Vorbereitung der Theoria motus unterblieb, bis er sie endlich im Jahre 1806 wieder vornahm. Im Jahre 1805 hatte er begonnen, die Störungen der Ceres zu entwickeln; aber auch hier gab es viel zu rechnen, und daß er schließlich selbst das Unbehagliche dieser dauernden mechanischen Beschäftigung empfand, zeigt ein Brief an Olbers vom 10. Mai 1805 in dem er sagt: „Die Methode, nach der ich die Ceresstörungen zu berechnen angefangen hatte, habe ich doch wieder aufgegeben. Das gar zu viele mechanische tote Rechnen, was ich dabei vor mir sah, hat mich abgeschreckt“. Auch die Nr. 122 des Tagebuchs gibt dem Ausdruck. In den Jahren 1806—1807 entstanden dann endlich die verfeinerten Methoden der Theoria motus.

10. Übersicht über die verschiedenen Methoden der Bahnbestimmung aus hypothetischen Werten von δ und δ'' oder von i und Ω.

Eine ausführlichere Schilderung der verschiedenen in betracht kommenden Methoden der Bahnbestimmung gibt Gauß in der Theoria motus, art. 124—129. Er bespricht dort zehn verschiedene Methoden, von denen die fünf ersten von hypothetischen Werten von δ, δ'' oder, was auf dasselbe hinauskommt, von r, r'' und die fünf letzten in ähnlicher Weise von hypothetischen Werthen für i und Ω ausgehen.

Bei der ersten Methode berechnet man aus δ und δ'' Radiusvektor, heliozentrische Länge und Breite im ersten und dritten Ort und hieraus i, Ω und die Längen in der Bahn v, v'', wie in der in Werke XI 1 abzudruckenden Notiz [3.] aus dem Handbuch Bb; hieraus findet man die Elemente wie in der Notiz [4.] a. a. O. und hieraus den mittleren geozentrischen Ort, wie in den Notizen [5.] und [6.] ebendort. Die Vergleichung mit dem beobachteten mittleren Ort liefert die beiden Daten (Fehler in Länge und Breite) für die Interpolation der verbesserten Ausgangswerte (vgl. Werke XI).

Bei der zweiten Methode rechnet man bis zu den Elementen wie in der vorigen; sodann leitet man aber jetzt den heliozentrischen Ort für die mittlere Beobachtung ab; den letzteren rechnet man außerdem aus dem beobachteten geozentrischen Ort und den für i und Ω gefundenen Werten. Die Vergleichung beider heliozentrischer Werte liefert die beiden Daten (Fehler in r und v) für die Interpolation.

Bei der dritten Methode rechnet man bis zur Bestimmung von i, Ω, v, v'', wie bei der ersten, sodann aus i, Ω und dem mittleren beobachteten geozentrischen Ort die Größen r' und v'; hierauf aus den drei heliozentrischen Örtern, also aus r, r', r'', v, v', v'' die Elemente und aus den Elementen die beiden Zwischenzeiten zwischen dem ersten und zweiten und zwischen dem zweiten und dritten Ort; die Vergleichung zwischen diesen errechneten Zwischenzeiten und den wahren geben hier die Daten zur Interpolation.

Die vierte Methode stimmt bis zur Auffindung von r, r', r'', v, v', v'' mit der dritten überein; hierauf rechnet man aber die Elemente einmal aus r, r', v, v' und zweitens aus r', r'', v', v''. Für jedes Element erhält man so eine Differenz zwischen beiden Systemen. Als Daten für die Interpolation wählt man zwei dieser Elemente aus.

Die fünfte Methode endlich verfährt wie die vierte; jedoch

führt man die Rechnung nicht bis zur Bestimmung der Elemente durch, sondern nur bis zu den Größen η und η'', die das Verhältnis zwischen Sektor und Dreieck ausdrücken (s. S. 34—35) und benutzt deren Differenz zur Interpolation.

Die fünf übrigen Methoden unterscheiden sich von den vorigen nur dadurch, daß man von hypothetischen Werten von i und Ω ausgeht; bei der sechsten und siebenten Methode rechnet man aus i, Ω und den beiden äußeren beobachteten Örtern r, r'', v, v'' und verfährt dann weiter, wie bei der ersten und zweiten Methode.

Bei der achten bis zehnten Methode leitet man aus i, Ω und allen drei beobachteten Örtern die Größen r, r', r'', v, v', v'' ab und verfährt dann weiter wie bei der dritten bis fünften. Für die siebente Methode gelten die Vorschriften der Notiz [7.] aus dem Handbuch Bb (Werke XI 1), und für die achte die der Notiz [8.] ebendort.

11. Vervollkommnete Methode der Theoria motus.

Bei der Schilderung dieser Methoden hält Gauß den Gesichtspunkt fest, daß man die ganze Rechnung für drei vorausgesetzte Wertepaare durchführt und es finden sich Beispiele für die siebente Methode in der in Band XI 1 der Werke abzudruckenden Notiz [III.] aus der Scheda Ah und für die achte in der Notiz [II.] aus Ag. Es handelt sich hierbei um das oben (S. 14) als „reine Versuchsmethode“ bezeichnete Verfahren.

Das Festhalten an dieser mag hier auffallen, da doch die vollkommenere in der Theoria motus wirklich durchgeführte Methode anstelle der Interpolation aus drei Wertepaaren die Methode der sukzessiven Annäherung setzt (vgl. Theoria motus, art. 135), die auf der folgenden Seite besprochen werden soll.

Oben ist geschildert, wie Gauß in der Theoria motus mit einem gewissen Anklang an seine ältesten Methoden sich einen Näherungswert für den Abstand r' des Himmelskörpers im mittleren Beobachtungsort verschafft (vgl. die Gleichungen 20 und 21). Dabei bilden hier die beiden Größen P und Q den Ausgangspunkt der Annäherungen, ähnlich wie es früher δ und δ'' oder i und Ω gewesen waren, und in der ersten Annäherung ist

$$28)\qquad P = \frac{\Theta''}{\Theta}, \quad Q = \Theta\Theta''.$$

Die den älteren Methoden vollkommen analog auf P und Q aufgebaute würde nun so vorgehen, daß man mit drei Hypothe-

sen, etwa: 1. $P_0 = \frac{\Theta''}{\Theta}$, $Q_0 = \Theta\Theta''$; 2. $P_1 = P_0 + \varepsilon_1$, $Q_1 = Q_0$; 3. $P_2 = P_0$, $Q_2 = Q_0 + \varepsilon_2$, drei Bahnen berechnet, welche die beiden äußeren Beobachtungen genau darstellen, und aus den Abweichungen der mittleren Beobachtung verbesserte Werte von P und Q interpolierte. Man sieht, daß der Nachteil der ältern Methoden darin liegt, daß man in allen drei Hypothesen die Rechnung bis zur Bestimmung der Elemente durchführen muß.

Die Methode der Theoria motus verläßt den Weg der versuchsweisen interpolatorischen Verbesserung der Ausgangswerte, indem sie erstens nur mit einer Hypothese P_0 und Q_0 rechnet und mit der hieraus gefundenen Bahn diese hypothetischen Werte unmittelbar verbessert und indem sie zweitens auch bei der Berechnung dieser ersten genäherten Bahn nicht bis zur Bestimmung der Elemente durchgeführt zu werden braucht.

Die strengen Werte von P und Q (Gleichung 19) lassen sich schreiben[1])

29) $$P = \frac{\Theta''}{\Theta} \cdot \frac{\eta}{\eta''}, \quad Q = \frac{r'^2}{rr''\eta\eta'' \cos f \cos f' \cos f''} \Theta\Theta''$$

und das Näherungsverfahren besteht darin, auf möglichst kurzem Wege aus den genäherten Werten von P und Q die Größen $\eta, \eta'', r, r', r'', f, f', f''$ zu finden, um die ersteren damit zu verbessern, bis alles stimmt. Erst wenn die schließlichen Werte dieser Größen gefunden sind, werden die Elemente berechnet. Aus den Gleichungen 24) und den vorhergehenden ist ersichtlich, wie r, r', r'' und f' gefunden werden; hier können auch gleich die Größen Θ und Θ'' wegen Aberration verbessert werden. Es folgt dann die Bestimmung von f und f'' aus den Gleichungen des art. 144:

30) $$\begin{aligned} \sin 2f &= r \sin 2f' \cdot \frac{n}{n'r'} \\ \sin 2f'' &= r'' \sin 2f' \cdot \frac{n''}{n'r'} \end{aligned}$$

und es sind jetzt η und η'' zu bestimmen.

Von der Bestimmung dieser Größen handeln die art. 88—94 der Theoria motus. Den Gang der äußerlich recht verwickelten

1) Es bedeutet: η und η'' das Verhältnis des Sektors zum Dreieck zwischen je zwei Orten,
f, f'' die halbe Differenz der wahren Anomalien:
$$f = \tfrac{1}{2}(v'' - v'), \quad f'' = \tfrac{1}{2}(v' - v).$$

Rechnung kann man durch die folgenden Formeln deutlich machen[1]):

$$31)\qquad m = \frac{\Theta}{\sqrt{8\cos^2 f.(rr')^{\frac{3}{2}}}},\quad 2l = \frac{\sqrt{\frac{r'}{r}}+\sqrt{\frac{r}{r'}}}{2\cos f} - 1$$

$$x = \sin^2\frac{g}{2}$$

$$32)\qquad \sqrt{l+x} + (l+x)^{\frac{3}{2}}\,\frac{2g-\sin 2g}{\sin^3 g} = m$$

$$33)\qquad \eta = \frac{m}{\sqrt{l+x}}.$$

Aus der Gleichung 32) ist g oder x zu bestimmen. Gauß entwickelt hierzu die Größe $X = \frac{2g-\sin 2g}{\sin^3 g}$ in eine Potenzreihe nach x und verwandelt diese in einen Kettenbruch. Er setzt sodann

$$X = \frac{\frac{4}{3}}{1-\frac{6}{5}(x-\xi)}$$

und erhält auf elementarem Wege den ersten der beiden für ξ (Werke VII, 1906, S. 117) angegebenen Kettenbrüche; vom zweiten sagt er, daß er seine Ableitung an anderer Stelle geben werde, da sie auf weniger elementaren Grundsätzen beruht. Dies ist in den „Disquisitiones generales circa seriem infinitam $1+\frac{\alpha.\beta}{1.\gamma}x+\cdots$" art. 14 (Werke III, S. 137) geschehen, wo er die Beziehungen zwischen den functiones contiguae und im besonderen die allgemeine Formel für die Kettenbruchentwicklung von $\frac{F(\alpha,\,\beta+1,\,\gamma+1,\,x)}{F(\alpha,\,\beta,\,\gamma,\,x)}$ benutzt. Die Vorarbeiten hierzu finden sich, wie diejenigen zu diesem Kapitel der Theoria motus überhaupt, im Handbuch Bd (Okt. 1805) und sind teilweise Werke X 1, S. 326 abgedruckt; man vergleiche auch die Bemerkungen von Schlesinger ebendort S. 330. Gauß entwirft eine Hilfstafel, die ξ als Funktion von x gibt und durch deren Benutzung die Bestimmung von x und hieraus die η von sich sehr einfach gestaltet.

In derselben Weise wird g'' bezw. x'' und daraus η'' bestimmt und somit findet man nach 29) genauere Werte für P und Q, mit denen die Rechnung (zweite Hypothese) wiederholt wird, bis man

---

[1]) $g = \frac{1}{2}(E''-E')$ bedeutet die halbe Differenz der exzentrischen Anomalien im zweiten und dritten Ort; ebenso wird $g'' = \frac{1}{2}(E'-E)$ gesetzt.

definitive Werte für η und η'' und für g und g'' hat; aus diesen ergeben sich in einfacher Weise (Th. m. art. 95, 96) die Elemente.

Die Vorarbeiten zu den art. 88—96 findet man im Handbuch Bb, S. 24—34, die nach der Tagebuchnotiz Nr. 125 „Methodum ex duobus locis heliocentricis corporis circa solem moventis eiusdem elementa determinandi novam perfectissimam deteximus" aus dem Januar 1806 zu stammen scheinen, und im Handbuch Bd, S. 37—47, die sich vermutlich auf die Tagebuchnotiz Nr. 126: „Methodum e tribus planetae locis geocentricis eius orbitam determinandi ad summum perfectionis gradum eveximus", vom Mai 1806, beziehen.

12. Besondere Fälle der Bahnbestimmung.

Die Bahnbestimmung eines Planeten aus den ersten nach der Entdeckung gemachten Beobachtungen kennzeichnet sich dadurch, daß der vom Planeten zurückgelegte Teil seiner Bahn und also die Winkel $2f$, $2f'$, $2f''$ zwischen den Radienvektoren klein sind.

Wenn einerseits naturgemäß eine Bahnbestimmung aus einem kleinen Bogen mit größerer Unsicherheit behaftet ist, so führt andererseits in diesem Falle die Gaußsche Methode hier sehr schnell zum Ziel, sodaß in der Regel schon die zweite Hypothese über P und Q zu ausreichend genauen Elementen führt (vgl. die Rechenbeispiele, art. 150—157 der Theoria motus); sie ist aber auch brauchbar, wenn die Beobachtungen weiter auseinander liegen; nur wird man, falls die Konvergenz der Hypothesen zu wünschen übrig läßt, sobald man mit der dritten Hypothese für P und Q die Rechnung bis zur Bestimmung des vierten Wertepaares dieser Größen durchgeführt hat, nicht mit diesem vierten Wertepaare weiterrechnen, sondern durch Interpolation aus den drei ersten Hypothesen einen besseren Näherungswert ansetzen (vgl. das Beispiel art. 158—161 der Theoria motus), ähnlich wie bei den älteren Methoden. Gauß hat von Anfang an Wert darauf gelegt, seinen Methoden allgemeine Giltigkeit zu geben und in der Theoria motus Beispiele gegeben, die die verschiedenen vorkommenden Fälle erleuchten. Er schreibt in einem in der Monatl. Corr. (Bd. VI, S. 275) abgedruckten Brief an v. Zach vom 8. Juli 1806: „Es ist mir übrigens überaus lieb, daß ich nicht schon 1802 meine Methode, wie ich die Ceres- und Pallasbahn berechnet hatte, bekannt gemacht habe, so viele Aufforderungen auch deshalb an mich gelangten. Denn seitdem habe ich noch immer an der Vervollkommnung der Methode selbst gearbeitet, besonders in dem vorigen Winter, und ihre jetzige Gestalt sieht ihrer ersten fast gar nicht mehr ähnlich. Um hiervon eine Probe zu geben, will ich nur eines

Umstandes erwähnen. Da das Problem so sehr verwickelt ist, so ist es der Natur der Sache nach nicht anders möglich, als daß bei der allerersten Annäherung einige Voraussetzungen gemacht werden müssen, die nur näherungsweise richtig sind (wie z. B. bei Dr. Olbers' Methode die ist, daß die Chorden bei der Erde und dem Kometen durch die mittleren radii vectores im Verhältnisse der Zwischenzeiten geschnitten werden). Voraussetzungen von dieser Art liegen also auch nach meiner Methode bei der ersten Annäherung zum Grunde, und zwar solche, die desto weniger von der Wahrheit abweichen, je kleiner die Zwischenzeiten sind. Man darf also die Beobachtungen, auf die man die erste Annäherung gründet, nicht gar zu weit von einander entfernt annehmen, weil man sonst vermöge der näherungsweise wahren Voraussetzung bei der ersten Rechnung gar zu weit von der Wahrheit zurückbleiben, und daher zu viele und beschwerliche Wiederholungen der Verbesserungs-Methoden machen müßte. Doch konnte ich, wie meine Methode 1802 war, bei der Pallas sogleich Beobachtungen anwenden, die 27 Tage auseinander waren; viel weiter hätte ich indeß doch nicht gehen mögen. Dagegen ist jetzt meine Methode so beschaffen, daß ich neulich, als ich die mir von Ihnen gütigst mitgeteilten Beobachtungen Orianis von 1805 zu einem für mein Werk bestimmten Exempel benutzen wollte, und also dieselben so behandeln mußte, als wenn ich von der Pallas-Bahn noch gar nichts wüßte, sogleich und zwar mit dem allerglücklichsten Erfolge die äußersten 71 Tage von einander entfernten Beobachtungen zum Grunde (vgl. Theoria motus art. 156) legen konnte, und es leidet gar keinen Zweifel, daß ich darin noch beträchtlich weiter hätte gehen können".

Im art. 159 der Theoria motus gibt Gauß ein Beispiel für die Bahnbestimmung der Ceres, bei dem die Beobachtungen 260 Tage auseinanderliegen.

Doch ist hierzu folgendes zu bemerken: In der Gleichung 12) sind $\frac{b}{a}$, $\frac{c}{a}$, $\frac{d}{a}$ bei ersten Bahnbestimmungen mit kurzen Zwischenzeiten große Zahlen zweiter Ordnung, und zwar ist b positiv, c und d negativ. Andererseits setzt Gauß die Größen $\frac{\vartheta}{r'^{3/2}}$, $\frac{\vartheta''}{r'^{3/2}}$, deren Summe $\frac{\vartheta'}{r'^{3/2}}$ und e als kleine Größen erster Ordnung voraus.

Im Juno-Beispiel ist rund

$$\frac{b}{a} = 78{,}0, \quad \frac{c}{a} = -38{,}4, \quad \frac{d}{a} = -108{,}5, \quad e = 0{,}25.$$

Beim Ceres-Beispiel finden dagegen andere Verhältnisse statt; hier ist:

$$\frac{b}{a} = 1{,}10, \quad \frac{c}{a} = +3{,}11, \quad \frac{d}{a} = -2{,}04, \quad e = 0{,}08$$

und

$$\frac{\vartheta}{r'^{3/2}} = 0{,}52, \quad \frac{\vartheta''}{r'^{3/2}} = 0{,}55, \quad \frac{\vartheta'}{r'^{3/2}} = 1{,}07.$$

Nimmt man die letztern als klein von erster Ordnung an, so wird $\frac{\vartheta''}{r'^{3/2}} - \frac{\vartheta}{r'^{3/2}} = 0{,}03$ klein mindestens von zweiter Ordnung.

δ' erscheint also nicht als Unterschied zweier großer Zahlen zweiter Ordnung, wie beim Juno- und ähnlichen Beispielen, bei denen Fehler in $\frac{n}{n'}$ und $\frac{n''}{n'}$ von großem Einfluß sind; obwohl diese Fehler hier an sich wegen der großen Zwischenzeiten größer sind, so werden sie durch die kleineren Faktoren $\frac{c}{a}$ und $\frac{d}{a}$ nicht merklich vergrößert[1]).

Sind die Zwischenzeiten außergewöhnlich groß, so wird auch immer aus den ersten Beobachtungen eine genäherte Bahn bereits bekannt sein und man wird dann nicht von den Näherungswerten 28) für P und Q ausgehen, sondern gleich genauere Werte für die erste Hypothese nach 29) anwenden (art. 163).

Es gibt bekanntlich eine Reihe von Fällen, in denen die Methode der Bahnbestimmung aus drei Beobachtungen versagt. Im art. 160—162 bespricht Gauß die Fälle, in denen die drei geozentrischen Örter so ausgewählt sind, daß sich die Bahn entweder garnicht oder nur sehr ungenau bestimmen läßt, nämlich 1. wenn die drei Himmelskörper Sonne, Erde, Planet bei einer der Beobachtungen ganz oder nahezu in einer geraden Linie sich befinden, oder wenn 2. der erste und dritte geozentrische Ort ganz oder nahezu zusammenfallen, oder endlich 3. wenn die drei geozentrischen Örter des Planeten und der heliozentrische Erdort in der mittleren Beobachtung ganz oder nahezu in einem größten Kreise liegen. Diese Fälle müssen durch eine andere Auswahl der benutzten Beobachtungen vermieden werden.

Den Fall, in dem zwei ganz verschiedene Bahnen den drei Beobachtungen genügen, bespricht Gauß im art. 142; die Entscheidung, in welcher von beiden der Himmelskörper sich bewegt, kann nur durch eine vierte entferntere Beobachtung erbracht werden.

1) Diese Bemerkung zum Ceres-Beispiel verdanke ich Herrn Frischauf.

Dem Falle endlich, in dem die Neigung der Bahn gleich Null oder sehr klein ist, paßt sich die Methode der Bahnbestimmung aus vier Beobachtungen (Theoria motus, art. 164—171) an, die im wesentlichen auf den gleichen Prinzipien beruht, wie die aus dreien. Gauß beabsichtigte anfangs, hier eine Bahnbestimmung des Uranus als Beispiel zu geben, konnte aber die dazu nötigen Beobachtungen nicht zusammenbekommen (vgl. den Briefwechsel zwischen Gauß und Olbers vom 27. und 31. Januar, 12. März 1807); er benutzte dann zu diesem Zweck die unmittelbar darauf entdeckte Vesta.

13. Rechnunng für parabelnahe Ellipsen und Hyperbeln; hyperbolische, parabolische und Kreisbahn-Bestimmung.

Im Tagebuch finden wir im April 1806 die Notiz Nr. 127: „Methodus nova ellipsin et hyperbolam ad parabolam reducendi". Diese bezieht sich auf die art. 30—46 der Theoria motus, in denen die Berechnung der wahren Anomalie aus der Zeit, und umgekehrt, für parabelnahe Ellipsen und Hyperbeln behandelt wird; entsprechende Vorarbeiten dazu stehen im Handbuch Bd, S. 52 f. und 91. Bemerkenswert sind dabei die scharfen Untersuchungen der art. 30—32 über die Genauigkeit, mit der sich die wahre Anomalie aus der mittleren ergibt; da diese Genauigkeit zu wünschen übrig läßt, wenn e nahe gleich Eins ist, so entstanden die Entwicklungen zur Reduktion parabelnaher Ellipsen und Hyperbeln auf die Parabel. Nach art. 30 scheint es, als ob Gauß an anderer Stelle sich weiter über die Genauigkeit logarithmischer Rechnungen überhaupt verbreiten wollte.

Zur Bestimmung einer hyperbolischen Bahn wendet Gauß ähnliche Methoden an, wie zu der einer elliptischen Bahn. In der Theoria motus gibt er jedoch ein Zahlenbeispiel nur für die Berechnung der Elemente aus zwei Radienvektoren, dem eingeschlossenen Winkel und der Zwischenzeit (art. 105).

Daß Gauß in der Theoria motus nicht auf die Bestimmung einer parabolischen Bahn eingeht, mag, wie schon oben gesagt, seinen Grund darin haben, daß diese Aufgabe schon von Olbers in brauchbarer Form gelöst war. Er begann erst später sich damit zu beschäftigen, als er selbst die Bahnen einiger Kometen berechnete. In einem Briefe an Olbers vom 3. Januar 1806 präzisiert er den von Olbers im § 37 seiner Abhandlung noch nicht klar

herausgeschälten Fall näher, in dem die Olbersschen Methode nicht anwendbar ist. Dies veranlaßte wahrscheinlich Olbers zu seiner Mitteilung vom Juni 1806 im Berliner astronomischen Jahrbuch für 1809 (W. Olbers, sein Leben und seine Werke, Bd. I, S. 76).

Im Jahre 1813 bei Gelegenheit der Beschäftigung mit dem Kometen 1813 II fand Gauß eine erhebliche Verbesserung der Olbersschen Methode; dort läßt sich in häufig vorkommenden Fällen die Berechnung der Radienvektoren der beiden äußeren Beobachtungen und der Sehne aus dem hypothetischen Abstand von der Erde im mittleren Ort nicht scharf genug ausführen. Gauß gibt hierfür andere Formeln, die den Hauptinhalt der „Observationes Cometae secundi a. 1813 . . .“ (Werke VI, S. 25) bilden und sich auch in seinem Briefe an Olbers vom 25. Juli 1813 finden.

Im Jahre 1815 endlich beschäftigte er sich eingehender mit der parabolischen Bahnbestimmung; er gibt nicht nur der Lambertschen Gleichung eine andere Gestalt, sondern entwickelt auch eine Reihe anderer wichtiger Beziehungen. Er begann den Entwurf einer Notiz „Allgemeine Theorie der Berechnung der Kometenbahnen“, dem er ein im Briefe an Olbers vom 13. Januar 1815 erwähntes Musterbeispiel zufügt. Die entsprechenden Bruchstücke aus dem Nachlaß sind in Werke VII, 1906, S. 323—374 abgedruckt; dazu gehört auch eine Tafel zur Berechnung der wahren Anomalie, die vollständiger ist als die Barkersche. Er schreibt darüber an Olbers am 29. Mai 1815: „Meine theoretischen Untersuchungen über die Berechnung der Kometenbahnen im allgemeinen hätte ich wohl einige Neigung, in Zukunft einmal in einem eigenen Werke bekannt zu machen, als Supplement zu meiner Th. M. C. C. Es könnte vielleicht 6—8 Bogen stark werden, und ich würde dann noch eine Tafel für die parabolische Bewegung von einer neuen Einrichtung beifügen, deren Gebrauch noch etwas bequemer ist als der der Barkerschen. Diese möchte auch noch 3 Bogen betragen. Der Grund dazu ist schon gelegt, wobei mir Herr Encke noch geholfen hat“ und an Schumacher am 5. Januar 1845: „Das Zusatzkapitel, welches ich einmal zu der Th. M. C. C. zu machen im Sinn hatte, würde die Berechnung der rein parabolischen Bahn betroffen haben. Eigentlich ausgearbeitet ist darüber niemals etwas gewesen; einzelnes ist aber so aufgeschrieben, daß ich es später danach wohl wiederherstellen könnte. Allein meine nächsten Arbeiten werden sich jedenfalls auf die Fortsetzung meiner geodätischen (theoretischen) Untersuchungen beziehen“.

Das Problem der Kreisbahn mag Gauß übergangen haben, weil sich nur in besonderen Fällen, z. B. wenn die Beobachtungen

zur Bestimmung einer Ellipse nicht ausreichen, die Berechnung einer solchen empfiehlt und diese „Arbeit sich durch eine sehr leichte und einfache Rechnung erledigen läßt" (Einleitung zur Theoria motus, Werke VII, 1906, S. 6); bei der Ceres hatte sich die Unzweckmäßigkeit der Berechnung einer Kreisbahn in vollem Maße gezeigt. Jedoch findet sich in dem bekannten Lehrbuch von Klinkerfues (s. Fußnote S. 21) eine Methode der Kreisbahnbestimmung, die von Gauß herrührt und im Bande XI 1 der Werke abgedruckt werden wird.

III. Bahnverbesserung aus einer grösseren Reihe von Beobachtungen.

1. Einleitendes.

Die Gaußschen Methoden zur Verbesserung einer Planetenbahn aus einer größeren Reihe von Beobachtungen treten uns an vielen zerstreuten Stellen seiner Veröffentlichungen und seines Nachlasses entgegen und daher dürfte eine Zusammenstellung erwünscht sein. Im dritten Abschnitt der Theoria motus finden wir in erster Linie die Begründung der zur Anwendung kommenden Methode der kleinsten Quadrate, in der Disquisitio de elementis ellipticis Palladis dagegen die eigentlichen astronomischen Formeln.

Über die Bildung der Normalörter spricht Gauß in den Artikeln 173—174 der Theoria motus; in seinen praktischen Rechnungen legt er die Normalörter meist in den Augenblick der Opposition in Länge, wodurch die Rechnung möglichst vereinfacht wird. Die Zeit der Opposition ergibt sich aus der Gleichheit der heliozentrischen Länge der Erde und der aus den Beobachtungen folgenden geozentrischen Länge des Planeten; beiden ist dann auch die heliozentrische Länge des Planeten gleich, die neben der aus den Beobachtungen folgenden geozentrischen Breite die Grundlage für die Bahnverbesserung bietet. Diese Methode ist natürlich nur dann anwendbar, wenn in der Nähe der Opposition Beobachtungen in genügender Zahl vorliegen[1]).

2. Bahnverbesserung aus vier Oppositionen.

In der Disquisitio de elementis ellipticis Palladis benutzt Gauß zur Bahnverbesserung zwei verschiedene Verfahren; das erstere (art. 5—8), das in ähnlicher Weise auch bei seinen Störungsrechnungen zur Anwendung kommt (vgl. S. 45—46), soll hier so ge-

1) Man vergleiche hierzu auch die Notizen Werke VII, 1906, S. 310.

schildert werden, wie es in der Disquisitio benutzt wird, indem nämlich nur vier Oppositionen benutzt werden; wegen des zweiten sehe man Seite 47. Beim ersteren findet die Methode der kleinsten Quadrate nur teilweise Anwendung; es ist das folgende:

Da Neigung und Knotenlänge bereits ziemlich genau bekannt zu sein pflegen, so ist es auch die Reduktion auf die Ekliptik und daher lassen sich die Längen in der Bahn, nur mit einem sehr kleinen Fehler behaftet, berechnen [1]). Man gewinnt hierdurch den Vorteil, daß man die Verbesserung der Elemente getrennt, einerseits für die Länge des Perihels, die Exzentricität, die mittlere tägliche Bewegung und die Länge in der Epoche, und andererseits für die Neigung und die Länge des Knotens ausführen kann; dadurch hat man nur mit Gleichungen mit vier Unbekannten zu operieren, von denen sich noch eine, die Verbesserung der Länge in der Epoche, sofort eliminieren läßt.

Aus den Längen in der Bahn v berechnet Gauß mit den genäherten Werten von π und φ die mittlere Anomalie M und ihre Ableitungen [2])

34) $$\frac{\partial M}{\partial \pi} = -\frac{\cos^2 \varphi}{(1 + e \cos w)^2} = m$$

und

35) $$\frac{\partial M}{\partial \varphi} = m \frac{(2 + e \cos w) \sin w}{\cos \varphi} = n$$

und erhält so für jeden Normalort eine Beziehung von der Form [2])

36) $$M = (M) + m \Delta\pi + n \Delta\varphi,$$

---

1) Man kann hierzu die direkte Formel

$$\operatorname{tg}(v - \Omega) = \frac{\operatorname{tg}(\alpha - \Omega)}{\cos i}$$

benutzen; schärfer rechnet man nach der Formel (vgl. Th. m. art. 50)

$$\sin(v - \alpha) = 2 \sin^2 \frac{i}{2} \sin(v - \Omega) \cos(\alpha - \Omega)$$

indem man für $v - \Omega$ den aus der ersteren berechneten Wert benutzt, wodurch man gleichzeitig eine Kontrolle erhält.

Es bedeutet: v die wahre Länge in der Bahn,
α die heliozentrische Länge.

2) M = mittlere Anomalie,
(M) = berechneter genäherter Wert der vorigen,
μ = mittlere Bewegung,
ε = mittlere Länge in der Epoche,
$t, t' \ldots$ = seit dem Periheldurchgang verflossene Zeit,
$w = v - \pi$ = wahre Anomalie.

also bei vier Oppositionen vier solche Gleichungen für M, M', M'', M'''. Da

$$M = \varepsilon - \pi + \mu t$$
$$M' = \varepsilon - \pi + \mu t'$$
u. s. w.,

so ergeben sich durch Elimination von $\varepsilon - \pi$ drei Gleichungen

$$\begin{aligned} \mu(t'-t) &= (M') - (M) + (m'-m)\Delta\pi + (n'-n)\Delta\varphi \\ \mu(t''-t') &= (M'') - (M') + (m''-m')\Delta\pi + (n''-n')\Delta\varphi \qquad (37)\\ \mu(t'''-t'') &= (M''') - (M'') + (m'''-m'')\Delta\pi + (n'''-n'')\Delta\varphi. \end{aligned}$$

Aus diesen läßt sich auch μ sofort eliminieren und daher $\Delta\pi$ und $\Delta\varphi$ durch eine kurze Rechnung bestimmen, worauf auch sogleich μ und ε gefunden werden.

Die auf diese Weise gefundenen Werte von μ, ε, π und φ sind noch mit dem Fehler der vorausgesetzten Werte von i und Ω behaftet, aber nur soweit dieser durch die Reduktion auf die Ekliptik eingeht, also mit einem sehr geringen Betrage. Man berechnet aus ihnen die Radiusvektoren in den vier Beobachtungen und erhält daraus die heliozentrischen Breiten γ nach der Formel[1]:

$$\sin(\beta - \gamma) = \frac{R \sin \beta}{r}. \qquad (38)$$

Um auf den Einfluß der Fehler in i und Ω Rücksicht zu nehmen, setzt Gauß

$$\gamma = (\gamma) + \left(\frac{d\gamma}{d\Omega}\right)\Delta\Omega + \left(\frac{d\gamma}{di}\right)\Delta i = (\gamma) + a\Delta\Omega + b\Delta i, \qquad (39)$$

wo (γ) der berechnete Wert für die heliozentrische Breite ist. Über die Differentialquotienten $\left(\frac{d\gamma}{d\Omega}\right)$ und $\left(\frac{d\gamma}{di}\right)$ sagt Gauß, daß er sie nicht auf analytischem Wege, sondern auf numerischem bestimmen wolle. Der Grund hierfür ist vielleicht der, daß die erstere Bestimmungsart, wenn sie streng ausgeführt wird, etwas verwickelt ist; denn $\left(\frac{d\gamma}{d\Omega}\right)$ und $\left(\frac{d\gamma}{di}\right)$ sind nicht gleich den partiellen Ableitungen von γ im gewöhnlichen Sinne, weil γ nicht aus den Elementen allein, sondern aus dem beobachteten β abgeleitet ist und

1) β = geozentrische Breite,
γ = heliozentrische Breite,
r = Radius vektor,
R = „ „ der Erde.

die Fehler Δi und $\Delta\Omega$ hier nur bei der Reduktion auf die Ekliptik und damit implizite in die Elemente π und φ, sowie auch in r eingehen. Gauß bestimmt $\left(\frac{d\gamma}{d\Omega}\right)$ und $\left(\frac{d\gamma}{di}\right)$, indem er die ganze vorausgehende Rechnung wiederholt unter Voraussetzung von etwas veränderten Ausgangswerten von i und Ω, sodaß er im ganzen drei Bestimmungen erhält; so wird bei ihm z. B. für die Opposition der Pallas[1]) von 1805

Hypothese	Ω	i	(γ)
I	172°28′46″.8	34°37′31″.5	−33°39′48″.15
II	172 29 46.8	34 37 31.5	−33 39 51.10
III	172 28 46.8	34 38 31.5	−33 39 35.63

und hieraus

$$\gamma = -33°39'48''.15 - 0.0492\,\Delta\Omega + 0.2087\,\Delta i$$

Er wendet also das Verfahren an, daß wir schon bei seinen ersten Bahnbestimmungen kennen gelernt haben und das wir heute als numerische Differentiation bezeichnen, wobei allerdings hier nur die ersten Differenzen berücksichtigt werden.

Drei weitere ähnliche Gleichungen ergeben sich für die übrigen Oppositionen.

Während die eben gefundenen Werte der heliozentrischen Breiten aus den beobachteten geozentrischen Breiten folgen, berechnet Gauß andererseits die heliozentrischen Breiten mit den (in Hypothese I.) vorausgesetzten Werten von i und Ω aus den beobachteten Längen nach der Formel[2])

40) $$\operatorname{tang}\delta = \operatorname{tang} i \sin(\alpha - \Omega)$$

und es wird

41) $$\delta = (\delta) + \frac{\partial\delta}{\partial\Omega}\Delta\Omega + \frac{\partial\delta}{\partial i}\Delta i = c\Delta\Omega + f\Delta i$$

wo

42) $$c = -\tfrac{1}{2}\sin 2\delta \operatorname{cotang}(\alpha - \Omega)$$
$$f = \frac{\sin 2\delta}{\sin 2i}.$$

1) Vgl. Werke VI, S. 12—13.

2) Hier bezeichnet:
(δ) die heliozentrische Breite, gerechnet nach der Formel 40),
δ den wahren Wert der vorigen,
α die heliozentrische Länge.

Für die Opposition von 1805 findet er z. B.

$$\delta = -33^{0}40'50.''63 + 0.1252\,\Delta\,\Omega - 0.9870\,\Delta i.$$

Da $\delta = \gamma$ sein muß, so liefert die Opposition von 1805 die Gleichung

$$43) \qquad 62.''48 - 0.1744\,\Delta\,\Omega + 1.1957\,\Delta i = 0.$$

Daneben ergeben die drei übrigen Oppositionen drei weitere ähnliche Gleichungen und da hiermit Δi und $\Delta\,\Omega$ überbestimmt sind, so leitet Gauß ihre Werte nach der Methode der kleinsten Quadrate ab.

Um nun endlich auch die vier übrigen oben gefundenen Elemente durch Beseitigung des Einflusses von Δi und $\Delta\,\Omega$ weiter zu verbessern, kann man entweder ihre genauen Werte durch Interpolation aus den drei Hypothesen für i und Ω bestimmen, oder, was zuverlässiger ist, die Bestimmung mit den scharfen Werten von i und Ω wiederholen (vgl. Disquisitio de elementis Palladis art. 8, Werke VI, S. 14—15).

3. Bahnverbesserung aus mehreren Oppositionen und mit Berücksichtigung der Störungen.

Legt man der Elementenverbesserung mehr als vier Oppositionen zugrunde, so erhält man mehr als vier Gleichungen der Form 37) und kann diese, wie die Gleichungen 43), nach der Methode der kleinsten Quadrate auflösen. Nimmt man auch auf die Störungen Rücksicht, so ist die Rechnung wesentlich die gleiche, nur muß man die gestörten Werte der Elemente benutzen. Das Verfahren, daß Gauß hier gewöhnlich anwandte, schließt sich an das vorerwähnte an. Es mag auch hier auseinandergesetzt werden; man vergleiche dazu Werke VII, 1906, S. 479—482, 486—488, 561—564.

Es seien i_0, Ω_0, π_0, φ_0 u. s. w. die wahren und (i_0), (Ω_0), (π_0), (φ_0) u. s. w. genäherte Werte der mittleren oder Normal-Elemente, wie Gauß sie nennt.

Indem man zu (i_0), (Ω_0), (π_0), (φ_0) die Störungen hinzufügt, erhält man die Näherungswerte für die gestörten (oskulierenden) Elemente, die mit (i), (Ω), (π), (φ) bezeichnet werden mögen. Mit den letzteren berechnet man wie im vorstehenden (Gleichung 36) die mittlere Anomalie (M) und hieraus die mittlere Länge $(L) = (M) + (\pi)$, so daß entsprechend den Gleichungen 36) der wahre Wert der gestörten mittleren Länge ist

$$44) \qquad L = (L) + (m+1)\,\Delta\pi + n\Delta\varphi.$$

Für die mittlere Länge gilt hier der Ausdruck

45) $$L = \varepsilon + \int \mu \, dt;$$

faßt man die Störungen dieser Größe als Störungen der mittleren Länge zusammen, indem man setzt (nach Werke VII, 1906, Seite 480)

$$\delta L = \delta \varepsilon + \delta \int \mu \, dt,$$

so hat man

46) $$L = \varepsilon_0 + \mu_0 t + \delta L = \varepsilon_0 + (\mu_0) t + \delta L + t \,.\, \varDelta \mu.$$

Zieht man also von (L) die Störungen δL und die aus dem Näherungswerte von μ folgende mittlere Bewegung seit der Epoche $t = 0$ ab, wobei man setzen kann

47) $$(\varepsilon_0) = (L) - \delta L - (\mu_0) t,$$

so erhält man für jede Opposition eine Gleichung der Form

48) $$\varepsilon_0 = (\varepsilon_0) + (m+1) \varDelta \pi + n \varDelta \varphi - t \varDelta \mu.$$

Gauß eliminiert aus diesen Gleichungen, deren Anzahl s sein mag, zunächst die Epoche ε_0, indem er das Mittel aus ihnen:

49) $$\varepsilon_0 = \frac{1}{s} \left\{ \Sigma(\varepsilon_0) + \varDelta \pi \,.\, \Sigma(m+1) + \varDelta \varphi \,.\, \Sigma n - \varDelta \mu \,.\, \Sigma t \right\}$$

nimmt und dieses von jeder der s Gleichungen abzieht; die entstehenden s Gleichungen für $\varDelta \pi$, $\varDelta \varphi$ und $\varDelta \mu$ löst er dann nach der Methode der kleinsten Quadrate.

Der weitere Verlauf der Rechnung zur Verbesserung der Werte von i und Ω ist derselbe wie im vorstehenden (S. 43); die Koeffizienten a und b können dabei in der Regel ihrer Kleinheit wegen fortgelassen werden, wodurch die Rechnung sich erheblich vereinfacht (vgl. Werke VII, 1906, S. 482).

Im Gegensatz zu dem soeben besprochenen Verfahren, das Gauß bei seinen Störungsrechnungen anwendet, finden wir zur Bahnverbesserung aus einer größeren Reihe von Oppositionen in der Disquisitio de elementis ellipticis Palladis (art. 10—14) ein direktes Verfahren. Gauß stellt hier die allgemein giltigen Gleichungen für die Differentialänderungen der heliozentrischen Länge und der geozentrischen Breite auf; nämlich [1])

---

1) l = heliozentrische Länge,
β = geozentrische Breite,
(l) und (β) genäherte Werte der vorigen,
b = heliozentrische Breite.

$$50)\quad l = (l) + \frac{a^2 \cos\varphi \cos i}{r^2 \cos^2 b}\Delta L + \frac{ta^2 \cos\varphi \cos i}{r^2 \cos^2 b}\Delta n$$
$$+ \left(\frac{\cos i}{\cos^2 b} - \frac{a^2 \cos\varphi \cos i}{r^2 \cos^2 b}\right)\Delta\pi$$
$$+ \frac{a^2 \cos i}{r^2 \cos^2 b}(2 - e\cos E - e^2)\sin E\,\Delta\varphi$$
$$+ \left(1 - \frac{\cos i}{\cos^2 b}\right)\Delta\Omega - \operatorname{tg} b \cos(l - \Omega)\,\Delta i,$$

$$51)\quad \beta = (\beta) - \frac{a \sin\beta \sin(\beta - b)\operatorname{tg}\varphi \sin w}{r \sin b}\Delta L$$
$$+ \left\{\frac{2\sin\beta\sin(\beta - b)}{3\mu \sin b} - \frac{a t \sin\beta \sin(\beta - b)\operatorname{tg}\varphi \sin w}{r \sin b}\right\}\Delta\mu$$
$$+ \frac{a \sin\beta \sin(\beta - b)\operatorname{tg}\varphi \sin w}{r \sin b}\Delta\pi$$
$$+ \frac{a \sin\beta \sin(\beta - b)\cos\varphi \cos w}{r \sin b}\Delta\varphi$$
$$+ \frac{2\sin\beta\cos(\beta - b)\cos b}{\sin 2i}\Delta i$$
$$- \sin\beta\cos(\beta - b)\cos b \operatorname{ctg}(l - \Omega)\,\Delta\Omega.$$

Da es sich um Oppositionen handelt, so kann man überall die geozentrische Länge statt l nehmen und daher die Gleichungen mit sechs Unbekannten direkt nach der Methode der kleinsten Quadrate lösen. Gauß bedient sich in der Disquisitio der Eliminationsmethode, die er dort ausführlich beschreibt.

Eine bequemere Form für die vorstehenden Gleichungen finden sich im Handbuch Be S. 49; sie sind Werke VII, 1906, S. 311—312 abgedruckt.

4. Allgemeine Formeln zur Bahnverbesserung.

Allgemeine Formeln zur Bahnverbesserung, bei denen nicht vorausgesetzt wird, daß die Beobachtungen sich auf Oppositionen beziehen, hat Gauß wohl nur bei seinen ersten Bahnbestimmungen angewandt; hier stellt sich die Rechnung weniger einfach. Gauß schildert zwei verschiedene hierfür in Betracht kommende Methoden in groben Zügen in der Theoria motus art. 187—189 und mit Berücksichtigung der Störungen ebendort art. 190—192. Die entsprechenden Differentialformeln gibt er in den Artikeln 76—77 der Theoria motus und im Handbuch Be Seite 56—57; die letzteren sind Werke VII, 1906, S. 296—297 abgedruckt; Gauß scheint von ihnen kaum Gebrauch gemacht zu haben.

Über die ausführliche Begründung der Methode der kleinsten Quadrate, die Gauß in der Theoria motus art. 175—186 gibt, wird an anderer Stelle zu berichten sein.

IV. Allgemeines über Gauss' störungstheoretische Arbeiten.

1. Einleitendes.

Für die Planeten Ceres und Pallas hat Gauß Störungen berechnet, die Werke VII, 1906, in größerem Umfange veröffentlicht worden sind. Während er die Pallas-Störungen nach einer mehr und mehr vervollkommneten Methode fast ganz durchgeführt hat, müssen die Berechnungen der Ceres-Störungen als Vorversuche bezeichnet werden; denn die hier angewandten Methoden waren noch nicht hinreichend entwickelt um — namentlich bei der Pallas — eine eingehendere Bestimmung der Störungen ohne fast unüberwindlich lange Rechnung zu ermöglichen. Während Gauß bei der Ceres Koordinatenstörungen anwendet, ist er nach längerem Schwanken, wofür er sich entscheiden solle, zuletzt bei der Pallas ganz zu Elementenstörungen übergegangen.

Er fand bei diesen Arbeiten Gelegenheit, seine mathematischen Untersuchungen anzuwenden, und zwar einerseits die über das arithmetisch-geometrische Mittel und die hypergeometrische Reihe, andererseits seine Interpolationstheorie. Er wird hiernach auf zwei grundsätzlich verschiedene Methoden geführt. Die erste beruht auf der analytischen Entwicklung der Störungsfunktion, wie sie auch schon von Laplace angewandt worden war: sie gibt die allgemeinen Ausdrücke für die Störungen analytisch dargestellt durch die Elemente, sodaß die Formeln für alle Planeten anwendbar sind, wenn die entsprechenden Zahlenwerte eingesetzt werden; Gauß benutzt dieses Verfahren bei der ersten Berechnung der Ceres-Störungen, bei der er sich auf die ersten Potenzen der Excentrizität beschränkt und Koordinatenstörungen rechnet. Bei der Weiterführung der Rechnungen erwies sich diese Methode als nicht durchführbar. Die zweite Methode beruht auf der interpolatorischen Entwicklung der Störungsfunktion nach der Theoria Interpolationis (Werke III, S. 265); sie ist die einzige in schwierigeren Fällen, wie sie z. B. die Pallas bietet, durchführbare. Man erhält dabei nicht die allgemeinen analytischen Ausdrücke, die für jeden Planeten gelten, sondern die Entwicklung wird mit Zahlenwerten entsprechend den Elementen eines bestimmten Planeten ausgeführt und gilt nur für diesen. Gauß hat es bei der Pallas benutzt. Für die zweite Berechnung der

Ceres-Störungen (Werke VII, 1906, S. 401 f.) bedient sich Gauß eines gemischten Verfahrens, indem er die Entwicklungen teils analytisch, teils interpolatorisch ausführt. Er hatte damals wohl die Vorzüge des interpolatorischen Verfahrens erkannt; er zögerte aber vielleicht, das analytische Verfahren ganz zu verlassen, weil er einen gewissen Wert auf die Anwendung seiner Untersuchungen über das arithmetisch-geometrische Mittel legte, die beim rein interpolatorischen Verfahren ganz in Wegfall kommt.

2. Geschichtliches über Gauß' Störungsrechnungen.

Gauß spricht sich Olbers gegenüber in einem Briefe vom 25. Juni 1802 dahin aus, daß er zwar die Störungen der beiden entdeckten Planeten berechnen wolle, daß er es aber für voreilig halte, die Störungen der im März 1802 entdeckten Pallas jetzt schon vorzunehmen, ehe sie eine längere Zeit hindurch beobachtet sei; er schreibt in diesem Briefe: „Auf die Störungen eher Rücksicht zu nehmen, ehe die Pallas einen weit größeren Bogen beschrieben hat, scheint mir übrigens ziemlich überflüssig. Eine ohne Rücksicht auf dieselben den Beobb. genau angepaßte Ellipse schließt diese schon mit ein und muß sie eine geraume Zeit einschließen. Es scheint mir daher nicht bloß eine unnötige, sondern selbst eine mißliche Arbeit, von Größen, deren kleine Unterschiede, und noch kleinere Unterschiede der Unterschiede, die Quelle unserer Kenntnis sein müssen, sehr große Quantitäten (wie die Störungsgleichungen gewiß bei Pallas sein müssen), die sich noch dazu bei dem jetzigen Zustande der Perturbationstheorie keineswegs sehr genau bestimmen lassen, erst abzuziehen, um sie nachher wieder hinzusetzen zu müssen. Wenn ich den Einfluß des Jupiter in diesem Jahre zu untersuchen Lust und Muße bekommen sollte, so würde ich einen anderen Weg einschlagen und die Elemente selbst als veränderlich ansehen; ich vermute sogar, daß dies bei der Pallas überhaupt auch künftig vorzuziehen sein wird; denn wenn man, wie sonst gewöhnlich, den nach mittleren Elementen berechneten Ort durch Gleichungen wird verbessern wollen, so denke ich, werden diese so groß und zahlreich sein, daß wenigstens meine Geduld schwerlich zureichen würde, eine große Anzahl Örter auf diesem Wege zu berechnen. Ich denke fast, es wird immer leichter sein, einen Ort aus Elementen ohne Tafeln zu berechnen, als mit Tafeln vielleicht 30 oder 40 Gleichungen für Länge, Breite und Radiusvector zu berechnen, und es scheint mir daher, daß es wohl vielleicht das Beste sein wird, daß man in den Tafeln wenigstens vor der Hand etwa von 3 Monat zu 3 Monat die veränderlichen oskulierenden rein ellip-

tischen Elemente angibt. Vielleicht wäre dies selbst bei der Ceres nicht ohne Nutzen. — Übrigens glaube ich auch, daß die Pallas nach einigen Umläufen das beste Mittel sein wird, die Masse des Jupiter zu bestimmen".

Schon in diesem Briefe deutet Gauß an, daß die Berechnung der Elementenstörungen vermutlich den Vorzug vor der der Koordinatenstörungen verdiene; doch scheint er in dieser Frage später unschlüssig gewesen zu sein (vergl. den folgenden Brief und S. 56—57), was darin seine Erklärung findet, daß es wesentlich von der angewandten Methode abhängt, welche Berechnungsart vorzuziehen ist. Bei seiner zuletzt bei der Pallas angewandten rein interpolatorischen Methode erwies sich in der Tat die Form der Elementenstörungen als die vorteilhaftere.

Am 12. Oktober 1802 schreibt er an Olbers: „Sollten Sie aber die ⚴ noch in der Hist. Cél. auffinden, so werde ich schwerlich der Begierde widerstehen können, davon noch Gebrauch zu machen. Doch bin ich selbst auf diesen Fall noch unschlüssig, ob ich die Elemente als veränderlich ansehen, oder die an die rein elliptisch berechneten Örter anzubringenden Störungen durch Quadraturen berechnen würde. Herr Burckhardt hat, wie ich aus der M. C. sehe, die Perturbationen bloß nach den ersten Potenzen der Exzentrizität und Neigung berechnet; dies Verfahren, welches bei der ⚳ vor der Hand mehrere Jahre hindurch völlig hinreichend sein wird, scheint mir bei der ⚴ wenig mehr zu helfen, als wenn man von den Störungen ganz abstrahiert. Überhaupt wird die Berechnung der Störungen der ⚴ künftig noch eine wahre Qual für die Analysten sein. Sich durch Quadraturen von einem Jahre zum anderen hinzuhelfen, hat freilich alle zu verlangende Schärfe, allein man wird doch auch wünschen, und die Würde der Wissenschaft fordert, daß man künftig, wenn man erst hinlängliche Beobb. hat, den Ort für jede entferntere Zeit bestimme. Allein wenn man die Störungen auf die gewöhnliche Art durch Reihen ausdrücken will, so werden diese äußerst langsam konvergieren, und ich glaube, daß man, wenn man den Ort jedesmal auf 1″ genau berechnen will, vielleicht mehrere hundert Gleichungen für Länge, Breite und Radiusvector nötig haben wird". Er sagt weiter in diesem Briefe, daß er für Ceres nach den VII. Elementen die Störungen berechnet, „aber bloß erst die erste Potenz der Exzentrizität in Betracht gezogen" habe.

Die letzten Worte beziehen sich auf die in Bd. VII, 1906, S. 377 f. abgedruckte erste Berechnung der Ceresstörungen, die er sehr schnell durchführte und deren Ergebnisse bereits im Dezember

1802 veröffentlicht wurden (M. C. Bd. VI; Werke VI, S. 227). Sie beruht im wesentlichen auf denselben Entwicklungen wie die Laplace'schen Untersuchungen. Die gewonnenen Ergebnisse brachte er in einer zur Anwendung sehr bequemen Form in Tafeln, welche ebenfalls (Bd. VI, S. 235) veröffentlicht sind. Die Rechnung sowie die Tafeln sind im folgenden (Seite 65 ff.) näher besprochen.

Gauß begnügte sich einstweilen mit diesen Ergebnissen. Er entwickelte sich nur, wie er im April 1803 an Olbers schreibt, die Differentialgleichungen für die Elementenstörungen, die er später anzuwenden gedachte; entsprechende Formeln finden sich in der Scheda Al, Seite 5.

Im Juni 1804 stellte die Pariser Akademie eine Preisaufgabe zur Berechnung der Störungen der kleinen Planeten, die sie in der „Gazette nationale ou le Moniteur universel“ Nr. 281, Samedi 11 messidor an 12 de la Republique (30. Juni 1804), S. 1276 mit folgenden Worten bekannt machte:

Sujet du prix de Mathématiques.

„Donner la théorie des perturbations de la planète Pallas, découverte par M. Olbers.

Les géomètres ont donné la théorie des perturbations avec une étendue et une exactitude suffisantes pour toutes les planètes anciennes connues, et pour toutes celles qu'on pourra découvrir encore, tant qu'elles seront renfermées dans le même zodiaque, et qu'elles n'auront qu'une excentricité peu considérable. Mercure était jusqu'à nos jours la plus excentrique de toutes les planètes, et en même temps celle qui avait l'inclination la plus forte; mais son peu de masse et sa position à l'une des limites du système planétaire la rendent peu propre à causer des altérations bien sensibles dans les mouvements des autres planètes. Uranus, découvert il y a vingt-trois ans par M. Herschell, se trouve placé à l'autre limite du système. Avec peu de masse et une excentricité médiocre, il a encore la plus petite de toutes les inclinations connues; en sorte que les formules qui avaient servi pour Jupiter et Saturne ont été plus que suffisantes pour cette planète moderne. Cérès, découverte il y a quatre ans par M. Piazzi, ayant avec une excentricité assez considérable, une inclination de 10° 38′, doit être sujette à de fortes et de nombreuses inégalités. Il paraît cependant que tous les astronomes qui ont travaillé à les déterminer se sont contentés des formules connues, dont le développement ne passe pas le produit de trois dimensions des inclinations et des excentricités. Ceux de cinq dimensions ont été employés dans la

Mécanique céleste pour un cas particulier, d'après une formule de M. Burckhardt. Le même astronome a présenté depuis à l'Institut national le développement général et complet des troisième, quatrième et cinquième ordres; mais ce degré de précision ne suffirait certainement pas pour la planète Pallas, dont l'excentricité est plus forte même que celle de Mercure, et l'inclination de 34° 37′, c'est à dire cinq fois plus grande que celle d'aucune autre planète connue. Il est même difficile de conjecturer quelles seront les puissances et quelles seront les dimensions des produits qu'il sera permis de négliger, et les calculs pourraient être d'une longueur, et les formules d'une complication telles qu'elles pourraient effrayer les astronomes les plus en état d'exécuter un pareil travail. Cette considération a déterminé la classe des sciences mathématiques et physiques de l'Institut national de France à proposer ce sujet pour le prix qu'elle doit distribuer dans sa séance publique du premier lundi de messidor an 14. En conséquence, elle invite les géomètres et les astronomes à discuter complètement toutes les inégalités de cette théorie, et à n'en omettre aucune qui ne soit reconnue entièrement négligeable; et comme ces inégalités pourraient varier assez sensiblement si les éléments elliptiques n'étaient pas encore assez exactement connus, il est indispensable que les concurrents ne se bornent pas à donner les coëfficients numériques des équations, ils doivent aussi donner les coëfficients analytiques, afin que l'on puisse y mettre les valeurs les plus exactes de la distance moyenne, de l'excentricité, de l'aphélie et de l'inclination, lorsque ces éléments seront mieux connus. Il résultera de ces coëfficients analytiques un autre avantage; ce que les planètes Cérès et Pallas, étant à des distances du Soleil si peu différentes, qu'il est même aujourd'hui très-difficile de dire laquelle des deux est la plus voisine ou la plus éloignée, la formule donnée pour Pallas pourra, sans beaucoup de changements, servir aussi pour Cérès, ainsi que pour toute autre planète qu'on pourrait découvrir par la suite, et dont on aurait de cette manière une théorie plus complète et plus certaine. La classe espère que la question paraîtra assez intéressante aux astronomes, pour qu'ils y donnent des soins proportionnés à la difficulté du sujet. Le prix sera une médaille d'or, d'un kilogramme.

Les ouvrages envoyés au concours devront être inscrits en français ou en latin, et ne seront reçus que jusqu'au 1er germinal an 14. Ce terme est de rigueur".

Gauß schreibt darüber an Olbers am 24. Juli 1804: „Daß das Nationalinstitut in Paris auf die Theorie der ⚴-Störungen einen Preis

gesetzt hat, weiß ich bloß vom Hörensagen, da ich Zeitungen nur unordentlich zu lesen pflege. Steht vielleicht im Moniteur oder sonstwo das Programm in extenso, so bitte ich um Nachweisung der Nummer. Mir deucht eine solche weitläufige Arbeit fast noch zu voreilig. Wenn indeß in diesem Jahre recht gute Meridianbeobb. gemacht und vom vorigen noch welche bekannt werden (ohne welche man die Neigung der Bahn und mehreres nur erst beiläufig bestimmen kann) — vielleicht werden sich in der Conn. des tems XIV., die ich noch nicht gesehen, noch welche von Bouvard, oder von dem Observat. de l'école militaire, von Burckhardt oder Le François finden —, so bin ich unter gewissen Umständen wohl geneigt, mich auf die Arbeit einzulassen“ und am 7. September desselben Jahres: „Das Programm im Moniteur wegen des Preises habe ich neulich gelesen. Einer solchen Art, die Störungen der Pallas zu behandeln, wie sie da verlangt wird, nämlich mit Hülfe analytischer Formeln, in denen man bloß die anderen Elemente eines andern Planeten, z. B. der ⚳, schlechtweg substituieren darf, um dessen Störungen zu finden, fühle ich meine Geduld nicht gewachsen; aber vornehmen werde ich nach meiner Manier die Pallasstörungen gewiß, obwohl wahrscheinlich die ⚳-Störungen früher. Denn auch bei der ⚳ scheint durch dasjenige, was ich bisher von den Störungen mitgenommen habe, noch wenig gewonnen zu sein. Etwas sonderbar scheint es mir, daß das Nationalinstitut selbst gesteht, daß die analytischen Formeln, wodurch es nun einmal das Problem gelöst wissen will, wohl so verwickelt sein möchten, daß sie auch den allergewichtigsten Astronomen abschrecken müssen“.

Im Januar 1805 bemerkte Gauß, daß seine Ceres-Störungen, namentlich die Sekularveränderung der Neigung und des Knotens, schlecht mit den Beobachtungen stimmten. Es war dies wohl die Folge eines Rechenfehlers; Gauß scheint sich nicht die Mühe genommen zu haben, diesen aufzusuchen. Er begann vielmehr, jetzt auch die Entwicklungen für die zweiten Potenzen der Exzentrizität nach der gleichen Methode auszuführen; diese seine Untersuchungen finden sich im Handbuch Bc, S. 50—54, und brechen auf der letzteren Seite gänzlich ab, da die Entwicklungen zu umständlich wurden; er schreibt am 10. Mai 1805 an Olbers: „Die Methode, nach der ich die ⚳-Störungen zu berechnen angefangen hatte, habe ich doch wieder aufgegeben. Das gar zu viele mechanische tote Rechnen, was ich dabei vor mir sah, hat mich abgeschreckt; auch selbst wenn alle Rechnungen, die ich Fremden hätte übertragen können, von Hrn. Bessel und Hrn. v. Lindenau (der sich gleichfalls mich bei dergl. Arbeiten zu unterstützen gefälligst erboten hat)

übernommen wären, würde für mich noch mehr übrig geblieben sein als meine Geduld hätte bestreiten können.

Ich habe indessen bereits eine andere Methode ausgesonnen, die eben so weit führen kann, als jene, aber bei weitem weniger — obwohl künstlichere — Arbeit erfordert. Ich habe schon stark angefangen, sie auf die Ceres anzuwenden, wiewohl vorerst nur nach einem eingeschränkteren Plane, indem ich nur bis an die 5. Potenzen der Exzentrizitäten von ♃ und ⚳ gehe. Diese Methode hat um so mehr Reiz für mich, da ich dabei von vielen, schon vor längerer Zeit angestellten, ziemlich tiefen Untersuchungen über eigene Arten von transzendenten Funktionen einen glücklichen Gebrauch machen kann. Ich werde in der Folge Ihnen eine Idee davon zu geben suchen. Auch hoffe ich, daß ich im Stande sein werde, alles so einzurichten, daß ich durch eine fremde Unterstützung eine ansehnliche Erleichterung erhalte, zwar nicht bei meiner diesmaligen Rechnung für die ⚳ (denn gerade in dem Teile, wo Hülfe zu brauchen wäre, bin ich schon selbst zu weit vorgerückt), aber doch wenn ich dieselbe wiederhole, welches nötig sein wird, da ohne Zweifel die erweiterten Störungsgleichungen noch mit ansehnlichen Änderungen in den Elementen selbst verbunden sein werden — oder auch wenn ich einst diese Arbeit bei der ⚴ und ⚵ vornehme, wo sie beträchtlich weitläufiger sein wird".

Die von Gauß hier erwähnte „andere Methode", die er „ausgesonnen", beruht einerseits auf seinen Untersuchungen über Interpolationstheorie, über die er an Olbers am 3. Januar 1806 schreibt, andererseits auf den Entwicklungen, die er im Briefe an Bessel vom 3. Sept. 1805 (Werke X 1, S. 237) auseinandersetzt. Es ist die oben (S. 49) erwähnte gemischte Methode, nach der die zweite Berechnung der Ceresbahn ausgeführt worden und die unten S. 77 f. besprochen ist. Gauß hat nach ihr nur die Breitenstörungen berechnet, welche am leichtesten durchzuführen waren, und gerade nach der ersten Rechnung nicht gestimmt hatten. Im weiteren erwies sich auch diese Methode ihrer Weitläufigkeit wegen nicht als durchführbar. Er schreibt an Olbers am 2. Juli 1805: „Mit meiner Rechnung der Störungen der ⚳ bin ich leider noch nicht fertig. Teils habe ich nicht immer anhaltend daran gearbeitet, teils habe ich auch daran bei weitem mehr Mühe gehabt, als ich vorher glaubte, hie und da auch wohl mehr, als notwendig gewesen wäre. Aber die zweckmäßigste Ausübung einer Methode lernt' man erst bei ihrer Anwendung. Mit den Störungen der Breite hatte ich angefangen. Ich hatte bei meiner ersten Rechnung alles, was von den Exzentrizitäten abhing, übergangen; Oriani hatte nur ein Glied mitgenommen. Nunmehr

habe ich alle Gleichungen, die ich über 1″ fand, mitgenommen, worunter auch ein paar sind, die von dem Produkte der Exzentrizitäten abhängen, also (da ohnehin in alle Breitengleichungen die Neigung entriert) von der Ordnung 3 sind. Von dieser Arbeit habe ich nun gewissermaaßen auch schon Früchte geerntet. Sie erinnern sich, daß ich mich schon seit 1803 beklagt habe, daß sich die Breiten in der ☍ von 1803 mit denen von 1801 nach der aus der Theorie gefundenen Bewegung des ☊ nicht mehr vereinigen ließen, und daß ich gezwungen war, den ♑ 1803 um 3′ weiter zu rücken. Ebenso stimmte in der ☍ 1804 die Breite nicht mit der von 1802; die Neigung mußte weit mehr verringert werden, als die Theorie angab.

Bei meinen neuesten Elementen mußte ich dem ☊ eine tägliche tropische Bewegung von 0″,241, und der Neigung eine tägl. Abnahme von 0″,0243 geben, um die Beobb. zu vereinigen, so daß

☊ 1801 Jan. 1	80° 54′ 46″	Neig.	1802 ☍	10° 38′ 1″
1803 ☍	80° 58′ 28″		1804 ☍	10° 37′ 38″.

Zu meiner großen Freude ist dieses nun nicht mehr nötig, und die Beobb. stimmen jetzt mit den neuen Breitengleichungen und Sekularbewegungen recht gut".

Die Fehler in den Breitenstörungen waren damit allerdings behoben. Gauß scheint nun aber nach allen Richtungen hin Versuche gemacht zu haben, die Methoden der Störungsrechnungen zu verbessern, und vor allem die ermüdenden langwierigen Entwicklungen und Rechnungen abzukürzen, nachdem er sich davon überzeugt hatte, daß die Laplace'sche Methode der Koordinatenstörungen, ebenso wie die von ihm selbst zunächst bei der Ceres angewandte, bei den kleinen Planeten nicht durchführbar ist. Zunächst versuchte er die Rechnung durch Entwerfung von Tafeln abzukürzen, die gleichzeitig auch für die anderen Planeten anwendbar waren.

Diese Tafeln beruhen auf der Entwicklung des Ausdruckes

$$(a^2 + a'^2 - 2aa' \cos\varphi)^{-\frac{1}{2}}$$

bei der ihn Bessel wesentlich unterstützte (vgl. unten S. 81); sodann aber ruhten seine Arbeiten über Störungstheorie fast vollständig bis zum August 1810. Er beklagt sich bei Olbers über das viele mechanische tote Rechnen und es mögen ihn in der Tat die langwierigen Rechnungen, die die von ihm benutzte Methode noch immer erforderte, von der Fortsetzung dieser Arbeiten abgeschreckt haben. Dazu kamen noch andere Gründe: seine im November

1807 erfolgte Übersiedlung nach Göttingen, der Tod seiner Frau und seines Sohnes, die Drucklegung der Theoria motus und vielleicht auch die traurigen politischen Verhältnisse; auch hat er sich in dieser Zeit vielfach nicht wohl gefühlt und klagt gegenüber Olbers (im Briefe vom April 1810) besonders über den unglücklichen Winter 1809/10. Die Weiterführung der Ceres-Störungen ist gänzlich unterblieben. Im Jahre 1810 beginnt er indessen, einen großen Teil seiner Zeit den Störungen der Pallas zu widmen.

Entsprechend seinen Äußerungen Olbers gegenüber (S. 49), daß die Berechnung der Störungen erst dann sich lohnen würde, wenn Pallas durch eine längere Reihe von Oppositionen hindurch beobachtet sei, untersucht er zunächst in der ersten Hälfte des Jahres 1810, wahrscheinlich im Frühsommer, ob sich bereits eine Einwirkung der Störungen nachweisen läßt; die entsprechenden Untersuchungen hat er in der Disquisitio de elementis ellipticis Palladis ... oppositionibus annorum 1803, 1804, 1805, 1807, 1808, 1809 (Werke VI, S. 1) veröffentlicht. Die Ergebnisse zeigen die starke Einwirkung der Störungen; die Differenzen in der Länge betragen bis zu $3\frac{1}{2}'$ und die in der Breite fast $1\frac{1}{2}'$.

Inzwischen hatte die Pariser Académie des Sciences den schon im Jahre 1804 ausgesetzten Preis auf die Berechnung der Störungen der Planeten immer von neuem aufgeschoben; Gauß schreibt an Olbers am 24. Okt. 1810: „Wüßte ich, daß das Institut die Preisfrage noch einmal prorogierte, so wäre ich nicht abgeneigt, die Pallas dazu zu wählen, sonst muß billig die Ceres den Vortritt haben, wo auch, weil bald die achte ☊ beobachtet wird, eine größere Satisfaktion zu erwarten ist. Bei der ⚴ würde ich doch wohl zuerst anfangen, die Elemente als variabel anzusehen und ihre Störungen während 7 Jahren durch Quadraturen zu bestimmen, wozu ich mir Formeln entworfen habe, die mir etwas bequemer scheinen als die La Placeschen. Falls es dann gelingt (wie es nicht anders zu erwarten ist), die 6 bisher beobachteten Oppositionen, die sich gar nicht mehr in eine Ellipse fügen wollen, gut zu vereinigen, so würde ich nach meiner schon vor 4 oder 5 Jahren entworfenen Methode die Störungen in der sonst üblichen Form, als periodische Störungen der Länge, der Breite und des Radiusvector, oder noch besser seines Logarithmen berechnen".

Danach beabsichtigte er zur Vorbereitung zuerst spezielle Störungen zu rechnen und zwar für die Elemente und sodann allgemeine Störungen für die Koordinaten. Auch hier zeigt sich, daß er zum Teil noch immer an den letzteren festhielt, obwohl

er schließlich ganz zu Elementenstörungen überging; denn er hat später auch bei den allgemeinen Störungen der Pallas Elementenstörungen benutzt.

Am 26. November 1810 berichtet er Olbers über seine Berechnung der speziellen Störungen und wiederholt, daß er vielleicht auch die allgemeine Theorie der Störungen unternehmen wolle. Er fragt, ob Olbers „durch seinen Wink in Paris eine nochmalige Verlängerung der Preisfrage veranlassen“ könne. Die hier erwähnte Berechnung der speziellen Störungen bezieht sich auf die im Band VII, S. 473—482 abgedruckte erste Rechnung (vgl. unten S. 83—84).

Am 13. Dezember 1810 schreibt Gauß an Olbers, daß er die Rechnung vollendet und eine seine „kühnste Erwartung übertreffende Übereinstimmung herausgebracht“ habe. Abgesehen von der schlecht beobachteten Breite der 5. Opposition (1808), die um 16″ abweicht, ist die größte Abweichung 6″. Andererseits zeigte sich, daß die letzte von Gauß auf Grund rein elliptischer Elemente abgeleitete und in der Monatl. Corr. veröffentlichte Ephemeride schon über 1° falsch war. Er erwähnt in diesem Brief auch, daß er angefangen habe, die Rechnung der speziellen Störungen nochmals schärfer (die erste Rechnung berücksichtigte nur die Störungen erster Ordnung) auszuführen; im Dezember wurde er hiermit fertig.

Diese zweite Rechnung ist in Bd. VII, 1906, S. 483 f. abgedruckt; sie ergab im wesentlichen eine ebenso gute Darstellung der Oppositionen, wie die erste; da schon die Ergebnisse der ersten Rechnung so befriedigend waren, konnte eine noch bessere Übereinstimmung kaum erwartet werden und die übrigbleibenden sehr kleinen Fehler sind auf die Einwirkung der übrigen Planeten zu setzen. Gauß berichtet über diese seine Rechnungen in der Monatl. Corr. Dezember 1810 und Jan. 1811 (Werke VI, S. 320—324) und sagt dort, daß er die Rechnung auch deswegen wiederholt habe, „weil bei dieser weitläufigen Arbeit hier und da Fehler sich eingeschlichen haben konnten“. Kleinere Fehler lassen sich in der Tat nachweisen und auch hieraus geht hervor, daß Gauß in der Regel auf eine genaue Prüfung seiner einzelnen Rechnungen verzichtete und es vorzog, selbst eine längere Rechnung ganz von neuem zu wiederholen.

Im Juli oder August 1811 begann Gauß die Berechnung der allgemeinen Störungen der Pallas. Er äußert sich jetzt über die Vorteile, die die Elementenstörungen bieten, in einem Briefe an Olbers vom 12. August 1811, nachdem er die Rechnung für

die Neigung und die Knotenlänge vollendet hatte: „Die 80 Gleichungen für Inkl. und ☊ ließen sich in 40 für die Breite zusammenziehen, ich glaube aber nicht, daß etwas gewonnen wird, denn wenn man die Elemente selbst stören läßt, so kann man ohne Bedenken einerlei gestörte Elemente als mehrere Monate gültig ansehen, und braucht also alle Jahre nur einmal für 6 Elemente die Störungen zu berechnen (vielleicht zusammen etwa 300—400 Gleichungen); dahingegen, wenn man bei den Elementen bloß Sekularänderungen anbringt und die periodischen bei Breite, Länge und Radiusvektor (zusammen vielleicht gegen 200), diese in einem Jahre doch wohl wenigstens für 6 verschiedene Örter berechnet werden müßten, um interpolieren zu können. Doch kann man dies in der Folge machen, wie man will, wenn nur erst alle Störungen in irgend einer Form da sind“.

Auch für die allgemeinen Störungen liegen zwei vollständige Rechnungen vor, welche beide gänzlich auf der interpolatorischen Methode beruhen.

Die erste Rechnung (vgl. unten S. 85), bei der die 5. Potenzen der Exzentrizitäten und die 23. Potenzen des Verhältnisses der halben großen Axen $\alpha = \frac{a}{a'}$ berücksichtigt sind, dürfte im März 1812 vollendet worden sein. Gauß verglich die Ergebnisse mit den Beobachtungen (den Oppositionen) und bestimmte die mittleren Elemente. Bei dieser Bestimmung scheint er die merkwürdige Entdeckung gemacht zu haben, daß die mittleren Bewegungen von Jupiter und Pallas sich wie 7 : 18 verhalten. Er schreibt darüber an Bessel am 5. Mai 1812: „. . . habe ich mich hauptsächlich mit den Pallasstörungen durch Jupiter beschäftigt. Sie werden darüber in Nr. 67 unserer Gelehrten Anzeigen einiges gelesen haben. Ihnen teile ich das merkwürdige daselbst in einer Chiffre niedergelegte Resultat gern mit, doch mit der Bitte, daß es vorerst ganz unter uns beiden bleibe. Es besteht darin, daß die mittlern Bewegungen von ♃ und ⚴ in dem rationalen Verhältnis 7 : 18 stehen, was sich durch die Einwirkung Jupiters immer genau wieder herstellt, wie die Rotationszeit unsers Mondes. Ich habe mit einer zweiten Berechnung der periodischen Störungen bereits einen Anfang gemacht . . .“.

Die gleiche Mitteilung hat er in einem verloren gegangenen Briefe an Olbers gemacht, wie aus Olbers Antwort[1]) vom 5. April 1812 zu ersehen ist; Bessel, und jedenfalls auch Olbers bittet er über diese Entdeckung vorläufig zu schweigen; er legt sie indessen

1) Werke VII, 1906, S. 421.

in den Göttingischen Gelehrten Anzeigen vom 25. April 1812 (Werke VI, S. 350) in der Chiffre

1111000100101001

nieder, zu der er „zu seiner Zeit den Schlüssel geben“ wollte. Das Letztere ist nun unterblieben und bei der Art und Weise, wie solche Chiffren gewählt zu werden pflegten, darf man auch nicht hoffen, die Lösung zu finden. Da man indessen aus dem Briefwechsel weiß, was sie bedeutet, so ist das auch gegenstandslos [1]).

Im Nachlaß haben sich nur vereinzelte ganz spärliche Notizen über diese Entdeckung auffinden lassen; sie sind in Bd. VII, 1906, S. 557—559 mit den nötigen Ergänzungen abgedruckt; vgl. unten S. 94.

Von vornherein bestand kein Zweifel, daß Gauß hier die Erscheinung meint, die Laplace bei den Jupitersmonden bereits entdeckt hatte (Méc. cél. Bd. IV 1805, Seite 16 und 64 f.) und die auch heut in der Theorie der kleinen Planeten eine Rolle spielt. Die Analogie mit der Rotation des Mondes, auf die Gauß hinweist, ist allerdings nur eine äußerliche; sie hat aber wohl schon Laplace veranlaßt, den Ausdruck Libration dafür zu gebrauchen.

Gauß schreckte nicht davor zurück, die ganze gewaltig umfangreiche Rechnung der allgemeinen Störungen zu wiederholen (vgl. unten S. 87) und dabei schärfer zu rechnen, indem er die 11. Potenzen der Exzentrizität berücksichtigte. Er begann diese Arbeit sofort im April (oder Ende März) 1812.

Über den ersten Teil dieser Rechnung hat Gauß ein ausführliches Tagebuch geführt, das vom 5. April bis zum 25. November 1812 reicht und angibt, welche Rechnungen er an einzelnen Tagen vollendet, wieviel Ziffern er gerechnet hatte, wieviel noch zu rechnen übrig waren, und wann diese voraussichtlich fertig sein würden. Dies Tagebuch ist in Bd. VII, 1906, S. 605—607 abgedruckt; mit „Präparation der Jupiter- und der Pallasörter“ bezeichnet Gauß die Berechnung der Koordinaten dieser Planeten, aus denen weiter die Komponenten T, V, W der Störungsfunktion folgen.

Bei der Berechnung der Koordinaten, die im Handbuch Bc, S. 97—99 steht, finden sich auch Notizen, nach denen Gauß in der Tat die Ziffern gezählt hat, die er beim Rechnen hat schreiben müssen. Eine Nachzählung der Ziffern würde schon an sich eine gewaltige Arbeit sein.

Fast das ganze Tagebuch handelt, bis 9. Juli, nur von der Berechnung und Interpolation der Größen T, V, W.

Nach dem Tagebuch erforderte diese das Schreiben von

1) Im dyadischen System gibt 111 die Zahl 7 und 10010 die Zahl 18.

338400 Ziffern; dagegen die Berechnung der Störungen der halben großen Axe 51040, und die für Knoten, Neigung und Exzentrizität zusammen circa 140000. Die ganze Arbeit wird man daher wohl auf 7—800000 Ziffern schätzen können. Gauß' tägliche Leistung bewegt sich zwischen 2600 und 4400 Ziffern und beträgt im Durchschnitt 3500. Man darf wohl annehmen, daß Gauß' Zeit zwischen dem 5. April und 25. November 1812 fast ausschließlich mit diesen Rechnungen ausgefüllt war. Dabei umfassen die schließlichen Störungsausdrücke einschließlich der sekularen Störungen:

im Knoten	105	Störungsglieder
in der Neigung	105	"
in der Perihellänge	184	"
im Exzentrizitätswinkel	124	"
im Logarithmus der halben großen Axe	145	"
in der mittleren Länge in der Epoche	161	"

Die zweite Rechnung der allgemeinen Jupiterstörungen der Pallas hatte Gauß im Juli 1813 vollendet; aber erst im Herbst 1816 begann er, deren Ergebnisse in Tafeln zu bringen, durch die die Berechnung der Störungsbeträge, die sonst einzeln aus den Störungsgleichungen berechnet werden müssen, außerordentlich erleichtert wird; diese Tafeln sind vollkommen erhalten und Werke VII, 1906, S. 572—577 abgedruckt. Ihre Berechnung wurde zum großen Teil von Encke ausgeführt und sie wurden im Spätherbst 1817 fertig.

Die Störungen der Pallas durch Saturn und Mars sind nicht bedeutend; schon im Frühjahr 1813 dachte Gauß daran, auch sie zu berücksichtigen. Die ersteren zu berechnen übernahm Nicolai in jener Zeit, und zwar nach einer Methode, die sich nur in unwesentlichen Punkten von der von Gauß bei den Jupiterstörungen benutzten unterscheidet. Sie wurden im Juli 1815 vollendet. In Gauß' Nachlaß finden sich die Briefe von Nicolai an Gauß nebst den Rechnungsergebnissen von Nicolai; die Rechnungen selbst werden auf der Heidelberger Sternwarte aufbewahrt. Von Gauß' Hand findet sich darüber nirgends etwas vor.

Wie Gauß verschiedentlich an Olbers berichtet (Briefe an Olbers vom 15. Juni 1814, 2. Dezember 1817), verschlechterte die Berücksichtigung der Saturnstörungen erheblich die sonst gute Übereinstimmung mit den Beobachtungen und es hat sich nicht aufklären lassen, ob sich bei den Nicolaischen Rechnungen ein Fehler eingeschlichen hat oder ob der Grund wo anders zu suchen ist. Es wäre auch heut noch von Interesse, diese Rechnungen

nochmals durchzusehen und womöglich zu veröffentlichen. Im Briefe vom 15. Juni 1814 sagt Gauß, daß die Übereinstimmung besser wird, wenn er allen Störungen der Epoche das entgegengesetzte Vorzeichen gibt. Da nach Werke VII, S. 587 bei den Störungen durch Mars die gleichen Formeln mit entgegengesetztem Vorzeichen benutzt sind, so liegt der Verdacht eines Fehlers im Vorzeichen allerdings nahe.

Die Berechnung der Störungen der Pallas durch Mars sollte nach einem Briefe von Gauß an Olbers vom 8. April 1813 (Bd. VII, 1906, S. 422) anscheinend Encke übernehmen; indessen sehen wir, daß Gauß selbst 1814 mit ihrer Berechnung beginnt und dabei teilweise eine neue besondere Methode anwendet, die man wieder, wie die zweite Rechnung der Ceresstörungen, als eine gemischt analytische und interpolatorische bezeichnen kann und bei der er das Verhalten der Funktion $r^2 + r'^2 - 2rr' \cos w$ im komplexen Gebiet untersucht (vgl. unten).

Schon im Sommer 1802 äußert Gauß in Briefen an Olbers sich dahin, daß die neuen Planeten und insbesondere Pallas ein vorzügliches Mittel bieten werden, um die Jupitermasse zu bestimmen und im April 1814 machte er den ersten Versuch hierzu, indem er auch diese in die Gleichungen zur Verbesserung der Elemente einführte; die entsprechende Rechnung ist in Bd. VII, 1906, S. 562 f. abgedruckt. Er fand hier den Wert $\frac{1}{1042,86}$. Nach einer späteren Rechnung, über die er im Briefe an Olbers vom 24. Juli 1816 berichtet und bei der die Saturn- und Marsstörungen berücksichtigt sind, fand er die Jupitermasse gleich $\frac{1}{1050}$, also einen sehr genauen Wert, da sie nach den besten Bestimmungen von Newcomb $\frac{1}{1047,4}$ beträgt.

Die Rechnung der Marsstörungen ist der einzige Teil von Gauß gewaltigem Unternehmen über die Pallas, der nicht vollendet ist und hier war die noch zu leistende Arbeit sehr gering.

Die Pariser Preisfrage war inzwischen auf das äußerste bis zum Jahre 1816 hinausgeschoben worden. Olbers schreibt an Gauß am 25. Januar 1815: „Ich hoffe doch . . . Sie werden Ihre Methode dem Pariser Institut mitteilen. Nur Ihretwegen hat man den Preis so lange offen gelassen". Gauß kam aber nicht dazu, die Arbeit fertigzustellen und einzusenden; er schreibt an Olbers am 8. Januar 1816: „Sollte der Preis nunmehro ausgegeben oder zurückgenommen sein, so würde ich meine Theorie nunmehro entweder stückweise in den Kommentaren oder auch in einer besonderen Schrift heraus-

zugeben denken". Nochmals scheint man Gauß Gelegenheit haben geben zu wollen, sich um den Preis zu bewerben; denn Olbers schreibt ihm am 7. März 1816: „... Jetzt eile ich nur, Ihnen wegen der Preisaufgabe zu Paris, wenn Sie es etwa noch nicht erfahren haben sollten, das Nähere zu melden. Der Ausspruch über den Preis wegen der Perturbationen der Planeten, namentlich der Pallas, ist wieder bis zum Jahr 1817 vertagt worden. Doch müssen die Schriften vor dem 1. Oktober 1816 eingesandt sein. Die Preisfrage war: „Die Theorie der Planeten, deren Exzentrizität und Neigung zu groß sind, als daß wir im Stande wären, ihre Störungen nach den schon bekannten Methoden genau zu berechnen". Die Klasse verlangt keine numerische Anwendung, sondern nur analytische Formeln, aber so eingerichtet, daß ein geschickter Rechner fähig sei, sie mit Sicherheit entweder auf den Planeten Pallas, oder auf einen andern der neu entdeckten oder noch zu entdeckenden Planeten anzuwenden. — Es waren nur zwei Abhandlungen eingelaufen, deren Verfasser aber die ausgesprochene Absicht der Klasse in der Preisankündigung nicht genug berücksichtigt haben. Beide (besonders der eine) haben noch zu mancherlei analytische Entwicklungen vorbeigelassen, die die Mathematiker noch erst machen müßten, um sie in den Stand zu setzen, die Auflösung des Problems, die sie gegeben haben, verstehen und beurteilen zu können. Sie haben es zu sehr versäumt, sich bis zu dem Standpunkte des Kalkulators herabzulassen, der nun wünschen sollte, Tafeln für die Pallas oder irgend einen andern Planeten zu bilden. Die Nachträge, die sie zu verschiedenen Zeiten eingeschickt haben, sind weit entfernt, alle diese Schwierigkeiten zu heben. Da die Klasse aus diesen Nachträgen und aus den eingeschickten Noten der anonymen Verfasser ersehen hat, daß sie nicht Zeit hatten, sich in alle die notwendigen Entwickelungen einzulassen, und zugleich in Erwägung zieht, daß auch vielleicht andere Mathematiker, die die Fähigkeit und Geschicklichkeit besitzen, diesen schwierigen Gegenstand zu behandeln, aus derselben Ursache abgehalten worden sind, als Preisbewerber aufzutreten, so hat sie die Preisausteilung noch bis Januar 1817 ausgesetzt. — Der Preis ist doppelt, eine goldene Medaille, 6000 Francs wert...".

Im oben erwähnten Brief vom 8. Januar 1816 sagt Gauß, daß er seine Theorie jetzt ausarbeiten wolle; wahrscheinlich war zu dieser Zeit die in Bd. VII, 1906, S. 439—472 abgedruckte Handschrift, von der zwei Entwürfe im Nachlaß vorhanden sind, zum Teil schon entstanden; man wird jedenfalls den zweiten Entwurf

in das Jahr 1816 versetzen müssen, während der erste auch schon vor diesem Brief, also 1815, abgefaßt sein könnte.

Da das Manuskript in französischer Sprache geschrieben ist, so war es wohl für die Einsendung nach Paris bestimmt; aber auch bis zu dem letzten von der Pariser Akademie gesetzten Termin, dem 1. Oktober 1816, wurde Gauß nicht fertig. Im Briefwechsel mit Olbers finden sich keine weiteren Anspielungen auf diese Angelegenheit; Gauß war in dieser Zeit sehr stark mit der Einrichtung der neuen Sternwarte beschäftigt und reiste im April zur Besprechung mit Reichenbach und Fraunhofer über instrumentelle Einrichtungen nach München; vom Juli 1816 bis Februar 1817 ruhte der Briefwechsel mit Olbers ganz. Jedoch setzte Gauß daneben zunächst noch seine Arbeiten über Pallas fort. In das Jahr 1817 fällt, wie oben erwähnt, die Entwerfung der Tafeln für die Jupiterstörungen. Im März 1818 hatte er die zwölf Oppositionen aus den Jahren 1803 bis 1817 verglichen und daraus neue mittlere Elemente bestimmt; er teilt dies Encke in einem Brief vom 25. März 1818 mit und bittet ihn, nun auch „die scharfe Vergleichung der sämtlichen Meridianbeobachtungen der Pallas vom April und Mai 1802 mit den osculierenden Elementen zu übernehmen".

Diese Rechnung hat Encke auch ausgeführt.

In der Folgezeit kam Gauß nicht wieder dazu, sich mit der Pallas oder überhaupt mit Fragen der theoretischen Astronomie eingehend zu beschäftigen. Neben den Untersuchungen und den Beobachtungen mit den neuen für die Sternwarte erworbenen Instrumenten war es vor allem seit dem Jahre 1820 die Hannoversche Landesvermessung, die ihn in Anspruch nahm. Er schreibt am 17. März 1822 an Schumacher: „... Was die Pallastafeln betrifft und die letzten Elemente, so sind alle darauf Bezug habenden Papiere so vereinzelt, daß es mir jetzt platterdings unmöglich ist, mich gleich wieder so hineinzustudieren, daß ich zur zuverlässigen Berechnung Anleitung geben könnte. Falls nicht noch etwas dazwischen kommt, was dieses Jahr die Fortsetzung meiner Messungen suspendiert oder verhindert, so müssen die Astronomen sich diesmal helfen so gut sie können ..." und am 8. Februar 1834 an Gerling „Ich habe dieser Tage (nach mehrjähriger Unterbrechung) die der Opposition nahe Pallas zu beobachten angefangen, wo Enckes Ephemeride etwa 5 Min. fehlt. Es ist mir dabei ein schmerzlicher Gedanke, daß meine vor mehr als 20 Jahren gemachte Arbeit über die Pallasstörungen ohne Fortsetzung, Entwickelung und Bekanntmachung bisher hat bleiben müssen, auch wahrscheinlich wie vieles Andere einst mit mir untergehen wird. Sie glauben nicht, wie

schwer es mir durch so vielfache Zersplitterung der Zeit so wie unter dem Druck so mancher Verhältnisse wird, eine wissenschaftliche Arbeit durchzuführen ...“ und endlich in einem Briefe an Bessel vom 21. März 1843 ...: „Die erste [Besselsche] Abhandlung über die Jupitermasse hat bei mir eine Erinnerung geweckt, die mir immer schmerzhaft ist, nämlich an meine alte Arbeit über die Pallasstörungen. Sie ist seit fast einem Vierteljahrhundert mir so fremd geworden, daß es mir schwer wird, mich selbst in den vorhandenen Papieren zu orientieren ...“; er fügt hinzu, daß „das letzte, was sich“ unter seinen Papieren „vorfindet, die Berechnung und Vergleichung der 14. Opposition vom 6. Januar 1820“ ist.

Bekanntlich fand Hansen bei der Bearbeitung der Störungen des Enckeschen Kometen ein ähnliches Verfahren zur Berechnung der Störungen wie das von Gauß angewandte, das er 1843 veröffentlichte; er schrieb darüber an Gauß und schickte ihm seine Abhandlung „Darlegung eines Verfahrens, um die absoluten Störungen der Himmelskörper, welche sich in Bahnen von beliebiger Neigung und elliptischer Exzentrizität bewegen, zu berechnen“ [1]). Gauß antwortet ihm am 11. März 1843: „Von einer Woche zur andern ist mein Dank für die gefällige Übersendung des Berliner Monatsberichts (der um dieselbe Zeit auch auf gewöhnlichem Wege mir zu Gesicht kam) verschoben, weil ich hoffte einige Zeit zu gewinnen, in den Gegenstand etwas weiter in meiner Antwort eingehen zu können. Leider ist diese Hoffnung getäuscht, und selbst in den bevorstehenden Ferien, für welche sich schon im Voraus so viele Rückstände und neue Abhaltungen gesammelt haben, darf ich mir kaum Hoffnung für einige freie Muße machen.

Ich beschränke mich daher darauf, meine Freude darüber auszusprechen, daß Sie bei den Perturbationsrechnungen auf ähnliche Art verfahren, wie ich schon vor mehr als 30 Jahren bei meinen weitumfassenden Rechnungen über die Pallasstörungen zu Werke gegangen bin, in so fern Sie den Gebrauch von Reihen nach den Exzentrizitäten und Neigungen ganz kassieren. Freilich haben Sie für die Kometenstörungen auch ganz besondere noch andere Untersuchungen nötig gehabt, zu denen für die Pallas keine Veranlassung sich fand und überhaupt in vielen andern Beziehungen abweichende Wege eingeschlagen. Bei den Störungen der Pallas durch Jupiter brauchte ich die Methode der variabeln Elemente und zwar mit Vorbedacht, denn obgleich man so 6 Elemente zu

---

1) Bericht der Kgl. Preußischen Akademie der Wissenschaften, Berlin 1843.

behandeln hat, während bei dem andern Verfahren nur halb so viele sind, so habe ich doch für den praktischen Gebrauch jenes vorgezogen; man braucht die Rechnung für jedes Jahr nur einmal (für einen Tag) zu machen, wozu ich eine besondere Hülfstafel konstruierte, vermittelst welcher in vergleichungsweise sehr kurzer Zeit die Rechnung absolviert werden kann, obgleich zusammen 801 Gleichungen (1602, wenn die Sinus- und Cosinus-Glieder desselben Arguments getrennt gezählt werden; es sind vollständig alle, deren Koeffizient über 0″,1 geht) berücksichtigt worden[1]). Durch sehr einfache Mittel kann man dann das Resultat für die ganze Beobachtungs-Saison ausreichend machen, ohne der Schärfe etwas zu vergeben. Auch die Störung durch Saturn wurde berechnet und die durch Mars nach einer wesentlich verschiedenen Methode angefangen, aber nicht vollendet. Andere immer weiter sich verzweigende Geschäfte haben mir später gar nicht erlaubt, auf jene Arbeiten wieder zurückzukommen, und es steht dahin, ob in meinem Alter ich Muße und Lust haben werde, mich wieder in die Sachen hineinzuarbeiten, da noch so viele andere Dinge sind, die ich eben so ungern untergehen lassen möchte. Sie sind sehr glücklich, daß Sie in einer äußern Lage sind, wo Sie Ihre Zeit nicht zu versplittern brauchen, und es wird mir jedenfalls zur Beruhigung gereichen, daß dieser Zweig der Astronomie bei Ihnen in den besten Händen ist. Ob ich bei den Marsstörungen etwas mit Ihrer mir angezeigten Integrationsart zusammenhängendes gebraucht habe, kann ich jetzt, wo mir die Sachen seit fast 25 Jahren entfremdet sind, nicht bestimmt ermitteln, möchte es aber fast bezweifeln . . .“.

V. Störungen der Ceres.

1. Erste Methode.

Im Oktober (oder schon im Sept.) 1802 machte sich Gauß an die erste Berechnung der Störungen der Ceres (Werke VII, 1906, S. 377 ff.), die er in der Form der Koordinatenstörungen ausführte, wesentlich nach den von Laplace und Schubert gegebenen Methoden, nach denen auch Oriani die Hauptstörungen der Ceres und Burckhardt die der Pallas berechnet hatte. Er entwickelte

1) Es ist wohl unnötig zu bemerken, daß die Arbeit, von der ich jetzt spreche, ganz verschieden ist von der gleichfalls von mir für den ganzen Zeitraum von 1802 bis etwa 1818 oder 1820 fortgeführten Rechnung durch Quadratur; diese nenne ich spezielle, jene generelle Rechnung und letztere hat in jener eine bei so ausgedehnten Rechnungen höchst notwendige Kontrolle gefunden.

sich aber seine Formeln selbst; die Entwicklungen finden sich im Nachlaß außerordentlich zerstreut in den Schedae und den Handbüchern vor.

Das ihm Eigne ist in erster Linie die Entwicklung der Störungsfunktion, bei der er seine Untersuchungen über das arithmetisch-geometrische Mittel ausnutzen konnte.

Die Gleichungen für die Störungen (Bd. VII, 1906 S. 381) nehmen bei ihm die folgende Form an[1]):

52)
$$\delta r = \frac{\mu a}{\sqrt{1-e^2}}\left\{\cos V \int r Q \sin V . n\,dt - \sin V \int r Q \cos V . n\,dt\right\}$$
$$\delta v = \frac{1}{\sqrt{1-e^2}}\left\{\frac{2 r\,d\varrho + \varrho\,dr}{a^2 n\,dt} + 3\mu a \iint \partial R . n\,dt + 2\mu a \int r \left(\frac{dR}{dr}\right) n\,dt\right\}$$
$$\Theta = \frac{\mu a}{\sqrt{1-e^2}}\left\{\cos V \int r \left(\frac{dR}{dz}\right) \sin V . n\,dt - \sin V \int r \left(\frac{dR}{dz}\right) \cos V . n\,dt)\right\}$$

wo[2])

53)
$$R = \frac{r \cos w}{r'^2} - \frac{1}{\mathfrak{R}}$$
$$\mathfrak{R} = \sqrt{r^2 + r'^2 - 2 r r' \cos w}$$
$$Q = 2 \int \partial R + r\left(\frac{dR}{dr}\right)$$
$$\left(\frac{dR}{dz}\right) = z'\left(\frac{1}{r'^3} - \frac{1}{\mathfrak{R}^3}\right)$$

Es kommt darauf an, alle unter den Integralzeichen vorkommenden Größen nach Vielfachen von nt oder der mittleren Längen und Anomalien von Ceres und Jupiter zu entwickeln, damit die

1) r, r' Radiusvektoren von Ceres und Jupiter
δr Störungen des Radiusvektor
δv „ der Länge in der Bahn
Θ Breite über der als fest angenommenen ungestörten Bahnebene
μ Jupitersmasse
a, e, n Bahnelemente der Ceres
V wahre Anomalie, wofür auch die wahre Länge v gesetzt werden kann
$\varrho = \delta r$

2) $z = r\Theta$ Abstand der Ceres von der ungestörten Bahnebene
z' Abstand des Jupiter von der ungestörten Bahnebene
$\mathfrak{R}$ Abstand zwischen Jupiter und Ceres
w Winkel zwischen den Radienvektoren von Ceres und Jupiter.
$\left(\frac{dR}{dr}\right)$ und $\left(\frac{dR}{dz}\right)$ partielle Ableitungen von R
∂R Ergebnis der Differentiation von R, wenn nur die Koordinaten der Ceres variiert werden.

Integrationen direkt ausgeführt werden können. Dies erfordert zunächst die Entwicklung von r, r', w, $\sin V$, $\cos V$ und sodann die von R und Q nach Vielfachen der mittleren Anomalien M und M' und damit auch nach Potenzen der Exzentrizitäten; diese Entwicklung führt Gauß direkt aus, indem er nur die ersten Potenzen der Exzentrizitäten berücksichtigt.

Ist nämlich etwa

$$\frac{1}{\mathfrak{R}} = f(r, r', w) \qquad f_0 = f(a, a', D) = (a^2 + a'^2 - 2aa' \cos D)^{-\frac{1}{2}},$$

wo $D = M - M' + \text{constans}$ und setzt man

$$\begin{aligned} r &= a(1 + \varDelta r) & \varDelta r &= e \cos M \\ r' &= a'(1 + \varDelta r') & \varDelta r' &= e' \cos M' \\ w &= D + \varDelta w & \varDelta w &= -2e \sin M + 2e' \sin M', \end{aligned}$$

wo $\varDelta r$, $\varDelta r'$, $\varDelta w$ die Exzentrizitäten enthalten, so ist

$$\frac{1}{\mathfrak{R}} = f_0 + \frac{\partial f_0}{\partial a} \cdot a \varDelta r + \frac{\partial f_0}{\partial a'} a' \varDelta r' + \frac{\partial f_0}{\partial D} \cdot \varDelta w.$$

Sodann sind aber f_0 und seine Ableitungen nach Vielfachen von D zu entwickeln.

Zur Entwicklung von f_0 braucht man die der Größe:

54) $$(a^2 + a'^2 - 2aa' \cos D)^{-\frac{1}{2}} = \tfrac{1}{2} P^0 + P' \cos D + P'' \cos 2D + \cdots$$

und zur Entwicklung von $\frac{\partial f_0}{\partial a}$, $\frac{\partial f_0}{\partial a'}$, $\frac{\partial f_0}{\partial D}$ die der Größe

55) $$(a^2 + a'^2 - 2aa' \cos D)^{-\frac{3}{2}} = \tfrac{1}{2} Q^0 + Q' \cos D + \cdots$$

Außerdem braucht man auch die Entwicklung von

$$(a^2 + a'^2 - 2aa' \cos D)^{-\frac{5}{2}}$$

um die Ausdrücke für ∂R und $\left(\frac{dR}{dr}\right)$ herzustellen.

Schon Laplace (Méc. cél., Bd. I., S. 267 ff.) und vor ihm Euler (Inst. calc. int. I § 279) zeigte, daß sämtliche Koefficienten dieser Reihen sich durch Rekursion berechnen lassen, sobald man P^0 und P' kennt, und auch Gauß leitet sich solche Rekursionsformeln ab. Sie sind hauptsächlich im Handbuch Bb, S. 16, entwickelt und finden sich auch an anderen Stellen des Nachlasses; abgedruckt sind sie Werke VII, 1906, S. 384. Zur Berechnung von P^0 und P' ist Laplace aber gezwungen, die bei nicht sehr kleinem a schwach konvergenten Reihen

56) $$\frac{1}{2}a' P^0 = 1 + \left(\frac{1}{2}\right)^2 \alpha^2 + \left(\frac{1.1}{2.4}\right)^2 \alpha^4 + \left(\frac{1.1.3}{2.4.6}\right)^2 \alpha^6 + \left(\frac{1.1.3.5}{2.4.6.8}\right)^2 \alpha^8 + \cdots$$
$$-\frac{a'^2}{a} P' = 1 - \frac{1.1}{2.4}\alpha^2 + \frac{1}{4}\cdot\frac{1.1.3}{2.4.6}\alpha^4 - \frac{1.3}{4.6}\cdot\frac{1.1.3.5}{2.4.6.8}\alpha^6 + \cdots$$

zu benutzen, wo $\alpha = \frac{a}{a'}$.

Gauß war infolge seiner Untersuchungen über das arithmetisch-geometrische Mittel im Stande, zur Berechnung dieser Koefficienten ein sehr stark konvergentes Verfahren einzuschlagen, da P^0 nichts anderes ist, als das reziproke arithm.-geom. Mittel aus $\frac{1}{2}(a+a')$ und $\frac{1}{2}(a-a')$ und auf ähnliche Weise auch P' gefunden werden kann. Ihm waren zu jener Zeit (1802) sicherlich die Integraldarstellungen

$$P^{(n)} = \frac{2}{a\pi}\alpha^{n+1}\int_0^\pi \frac{\sin^{2n}\varphi\, d\varphi}{\sqrt{1-\alpha^2\sin^2\varphi}}$$

geläufig und ebenso die Beziehungen dieser Integrale für $n = 0$ und 1 zum arithmetisch-geometrischen Mittel. In der unvollendeten Handschrift „De origine proprietatibusque generalibus numerorum mediorum arithm.-geometricorum“ (Werke III., Seite 361 f.), in der er sich ausdrücklich auf die Anwendung seiner Untersuchungen zur Entwicklung der Störungsfunktion bezieht, befolgt er zur Berechnung von P^0 folgenden Gedankengang[1]):

Es sei

57) $$\frac{1}{\sqrt{1-u^2\cos^2\varphi}} = P + 2Q\cos 2\varphi + 2R\cos 4\varphi + 2S\cos 6\varphi + \cdots$$

Indem man auf beiden Seiten zwischen den Grenzen 0 und π integriert, erhält man

58) $$\int_0^\pi \frac{d\varphi}{\sqrt{1-u^2\cos^2\varphi}} = \pi P.$$

Andererseits ist aber

59) $$\int_0^\pi \frac{d\varphi}{\sqrt{1-u^2\cos^2\varphi}} = \int_0^\pi \left\{1 + \frac{1}{2}u^2\cos^2\varphi + \frac{1.3}{2.4}u^4\cos^4\varphi + \cdots\right\} d\varphi$$
$$= \left\{1 + \left(\frac{1}{2}\right)^2 u^2 + \left(\frac{1.3}{2.4}\right)^2 u^4 + \cdots\right\}\pi$$

1) Vergl. auch Schlesinger, Materialien III. Seite 53 f.

Es ist also

60) $$P = 1 + \left(\frac{1}{2}\right)^2 u^2 + \left(\frac{1.3}{2.4}\right)^2 u^4 + \cdots$$

In der Handschrift wird nun bewiesen, daß die Entwicklung des reziproken arithm.-geometrischen Mittels aus $1+u$ und $1-u$ nach Potenzen von u ebenfalls die vorstehende Reihe ergibt; es ist also

61) $$P = \frac{1}{M(1+u, 1-u)} = \frac{1}{M(1, \sqrt{1-u^2})}.$$

Der Ausdruck $(a^2 + a'^2 - 2aa' \cos D)^{-1/2}$ läßt sich auf die Form $\frac{1}{c\sqrt{1-u^2\cos^2\varphi}}$ bringen, wenn man setzt

62) $$u^2 = \frac{4aa'}{(a+a')^2}, \quad c = a + a', \quad \varphi = \frac{D}{2}.$$

Es ist also

63) $$P^0 = \frac{2}{c} P = \frac{1}{M\left(\frac{c}{2}, \frac{c}{2}\sqrt{1-u^2}\right)} = \frac{1}{M\left(\frac{a'+a}{2}, \frac{a'-a}{2}\right)}.$$

Zur Bestimmung von Q und damit auch von $P' = \frac{2}{c} Q$ bietet der zweite Teil der Handschrift nur die Entwicklung von $dM(x, y)$, mit der sie abbricht; da Gauß hier ein Zahlenbeispiel vom Planeten Ceres entnimmt, so ist wohl anzunehmen, daß hier auch die Beziehungen zur Bestimmung von Q abgeleitet werden sollten. Obwohl uns die Untersuchungen über das arithm.-geometrische Mittel an den verschiedensten Stellen des Nachlasses (vergl. den Aufsatz von Schlesinger, Materialien Nr. II und Werke XI 1, S. 172) entgegentreten, so findet sich doch an keiner dieser Stellen die direkte Ableitung des Ausdruckes für P' in der Form, in der er hier gebraucht wird. Doch kann man im Hinblick auf das vorige wohl auf folgenden Gedankengang zur Ergänzung der genannten Handschrift schließen: Man hat

$$\frac{\cos 2\varphi}{\sqrt{1-u^2\cos^2\varphi}} = Q + (P+R)\cos 2\varphi + (Q+S)\cos 4\varphi + \cdots$$

und hieraus

$$\int_0^\pi \frac{\cos 2\varphi \, d\varphi}{\sqrt{1-u^2\cos^2\varphi}} = \pi Q,$$

also

$$64)\qquad \int_0^{\pi} \frac{\cos^2\varphi\, d\varphi}{\sqrt{1-u^2\cos^2\varphi}} = \frac{\pi}{2}(P+Q).$$

Entwickelt man diesen Ausdruck ähnlich wie den Ausdruck 59), so erhält man

$$65)\qquad \frac{1}{2}(P+Q) = \frac{1}{2} + \left(\frac{1}{2}\right)^2 \cdot \frac{3}{4}u^2 + \left(\frac{1.3}{2.4}\right)^2 \frac{5}{6}u^4 + \left(\frac{1.3.5}{2.4.6}\right)^2 \cdot \frac{7}{8}u^6 + \cdots;$$

hieraus und aus 60) folgt

$$66)\qquad \frac{1}{2}(P-Q) = \frac{1}{2} + \left(\frac{1}{2}\right)^2 \cdot \frac{1}{4}u^2 + \left(\frac{1.3}{2.4}\right)^2 \cdot \frac{1}{6}u^4 + \left(\frac{1.3.5}{2.4.6}\right)^2 \cdot \frac{1}{8}u^6 + \cdots$$

Aus den beiden letzten Entwicklungen kann man auch die für Q herstellen, worauf wir jedoch verzichten.

Differenziert man 60) nach u^2, so wird

$$67)\qquad \frac{dP}{du^2} = \left(\frac{1}{2}\right)^2 + \left(\frac{1.3}{2.4}\right)^2 2u^2 + \left(\frac{1.3.5}{2.4.6}\right)^2 3u^4 + \cdots$$

und hieraus findet man

$$68)\qquad (1-u^2)\frac{dP}{du^2} = \frac{1}{4} + \left(\frac{1}{2}\right)^2 \frac{1}{8}u^2 + \left(\frac{1.3}{2.4}\right)^2 \frac{1}{12}u^4 + \left(\frac{1.3.5}{2.4.6}\right)^2 \frac{1}{16}u^6 + \cdots$$

also nach 66)

$$(1-u^2)\frac{dP}{du^2} = \frac{1}{4}(P-Q)$$ [1]

und

$$69)\qquad \frac{Q}{P} = 1 - 2\frac{1-u^2}{u}\frac{1}{P}\frac{dP}{du}.$$

Nach 61) ist aber

$$70)\qquad \frac{1}{P}\frac{dP}{du} = -\frac{1}{M(1,\sqrt{1-u^2})}\frac{dM(1,\sqrt{1-u^2})}{du}.$$

Am Schlusse der genannten Handschrift (Werke III, Seite 373) gibt Gauß die Formel

$$71)\qquad dM(x,y) = \frac{M(x,y)}{2z}\left\{\frac{dx}{x}(z+2z'+4z''+\cdots) + \frac{dy}{y}(z-2z'-4z''\cdots)\right\}$$

wo

$$72)\qquad \begin{aligned} z &= x^2-y^2 \\ z' &= x'^2-y'^2 \\ z'' &= x''^2-y''^2 \\ &\cdots\cdots \end{aligned}$$

---

1) Man vergl. hierzu die Bemerkungen von Schlesinger, Werke X 1, S. 265.

und wo

$$x' = \frac{x+y}{2},\; y' = \sqrt{xy},\; x'' = \frac{x'+y'}{2},\; y'' = \sqrt{x'y'},$$

die nach dem Algorithmus des arithm.-geometrischen Mittels gebildeten Größen sind. Hieraus hat man

$$73)\quad \frac{1}{M(1,\sqrt{1-u^2})}\,\frac{dM(1,\sqrt{1-u^2})}{du} = -\frac{1}{2z}\,\frac{u}{1-u^2}(z-2z'-4z''-\cdots)$$

und daher nach 69) und 70)

$$74)\quad \frac{Q}{P} = \frac{2}{z}(z'+2z''+4z'''+\cdots);$$

die Größen $z, z', z'' \ldots$ sind nach 72) zu bilden, und zwar, da sie mit einer willkürlichen Konstanten multipliziert werden können, mit den folgenden Werten

$$x = C,\quad y = C\sqrt{1-u^2}.$$

Wenden wir die Formel 74) auf die Entwicklung 54) an, so wird in Uebereinstimmung mit der Gleichung für P' in Werke VII, 1906, Seite 385

$$75)\quad \frac{P'}{P^0} = \frac{2}{z}(z'+2z''+3z'''+\cdots)$$

wo nach 62)

$$x = C,\quad y = C\,\frac{a'-a}{a'+a}.$$

Wählt man

$$C = \frac{a'+a}{2}$$

so erhält man auch für $z, z', z'', \ldots$ die gleichen Werte wie in Werke VII.

In dem Werke X 1, Seite 237, abgedruckten Briefe an Bessel gibt Gauß eine später von ihm aufgefundene Entwicklungsmethode, über die man unten Seite 81 vergleiche. Eine Kettenbruchentwicklung für den Quotienten $\frac{P'}{P^0}$ findet sich jedoch schon 1800 in der Scheda Ac; sie ist abgedruckt Werke X 1, Seite 185; man vergleiche die Bemerkungen von Schlesinger, ebenda, Seite 268.

Die Integrationen bieten, sobald die Entwicklungen ausgeführt sind, nichts besonderes, da nur die Divisoren anzubringen sind;

die sekularen Glieder verwandelt Gauß in Störungen der Exzentrizität und der Aphellänge bezw. der Neigung und der Knotenlänge.

Gauß stützte seine Rechnung auf die VII. Cereselemente; mit Hilfe der so gefundenen Störungen verbesserte er diese Elemente und berechnete die Störungen aufs neue, um daraus schließlich die VIII. Elemente abzuleiten. Mitte Oktober 1802 hatte er diese Arbeit vollendet, wie er Olbers in einem Briefe vom 12. Okt. dieses Jahres mitteilt; die Störungsgleichungen sind in der M. C. Bd. VI. Seite 492—498 (Werke VI. S. 227—230) veröffentlicht.

2. Tafeln der Ceresstörungen.

Zur Vergleichung der Beobachtungen mit der Rechnung und zu der darauf beruhenden Verbesserung der Elemente ist der Betrag der Störungen nach den erhaltenen Störungsgleichungen zu berechnen. Diese langwierige Arbeit kann durch Entwerfung von Tafeln abgekürzt werden.

Gauß[1]) sagt darüber: „In der Zeit, daß ich nach den Störungsformeln den numerischen Wert berechne, könnte ich mehr wie eine Bahnbestimmung machen“ und entwirft nach folgenden Grundsätzen eine Tafel der Ceresstörungen:

Die von den Exzentrizitäten unabhängigen Störungsglieder hängen nur vom Argument ⚳ − ♃[2]) und seinen Vielfachen ab, können also für dies Argument tabuliert werden. Diesen Teil der Störungen nennt Gauß den ersten Teil und tabuliert ihn (Werke VI. S. 237, vgl. auch Werke VII, 1906 S. 396 f.) in Tafel I (Länge) und Tafel III (Radiusvektor). Die Störungsglieder, die die erste Potenz der Exzentrizitäten enthalten, sind von der Form

$$\alpha_i \frac{\sin}{\cos} (iD - ⚳ - c_i)$$

wo α_i und c_i numerisch gegebene Konstanten sind, $D =$ ⚳ − ♃ ist und i eine ganze positive oder negative Zahl oder Null bedeutet. Die Summe aller dieser Glieder mit verschiedenem i läßt sich auf die Form bringen (vgl. Bd. VI. S. 236)

$$A \sin (\varphi - ⚳),$$

wo A und φ nur von D abhängen und für dieses Argument tabuliert werden können. Hierbei gebraucht Gauß einen beson-

1) Brief an Olbers vom 12. Okt. 1802.

2) ⚳ mittlere Länge der Ceres
♃ „ „ des Jupiter.

deren Kunstgriff; man kann nämlich allgemein statt φ auch eine Größe $B+nD$ einführen, wo dann B gleichfalls als Funktion von D tabuliert werden kann und die obige Summe die Form

$$A \sin(B+nD-♀)$$

annimmt. Der Vorteil der Einführung von B liegt darin, daß man n so wählen kann, daß B sich möglichst langsam mit der Zeit, also mit D, ändert, wodurch die Entnahme aus der Tafel wesentlich erleichtert wird. Gauß wählt für Ceres $n=2$, offenbar weil das Hauptstörungsglied das Argument $2D-♀-c_2$ hat.

Auffallend ist, daß Gauß auch die kleinsten Störungsglieder berücksichtigt und nur beiläufig (Bd. VI. S. 235) von der Fortlassung der kleineren, „die einzeln unter 2″ betragen", spricht, und daß er die Tafelwerte bis auf Zehntel Sekunden angibt, obwohl er doch bereits die von den zweiten Potenzen der Exzentrizitäten abhängenden Glieder nicht berücksichtigt hat.

Auch hier zeigt sich, wie bei vielen anderen Gelegenheiten, Gauß' Neigung die Rechnung mit weit mehr Stellen auszuführen, als der Genauigkeit der zugrundeliegenden Zahlen entspricht.

3. Gauß' zweite Methode zur Berechnung der Störungen der Ceres. (Theoria interpolationis).

Da Gauß von den Ergebnissen seiner ersten Methode nicht befriedigt war, schlug er einen neuen Weg zur Entwicklung der Störungsfunktion und der davon abhängenden Größen ein.

Ganz neue Gesichtspunkte boten sich ihm hier durch seine Untersuchungen über Interpolationstheorie, wie sie in seiner nachgelassenen Handschrift „Theoria interpolationis methodo nova tractata" niedergelegt sind (Werke III. 265). Es scheint, daß dies Manuskript zum Teil unter der Einwirkung des Gedankens der Anwendung auf die Berechnung der Störungen entstanden ist; insbesondere kommen die Untersuchungen der art. 10f. in Betracht.

Es tritt hier an die Stelle der analytischen Entwicklung einer periodischen Funktion in eine Reihe, bei der die Herstellung der allgemeinen Ausdrücke für die Koeffizienten allzu verwickelt wird, die interpolatorische, rein numerische Entwicklung: Der numerische Wert der zu entwickelnden Funktion wird für eine Reihe von Werten des Argumentes berechnet, die regelmäßig auf den ganzen Umkreis von 2π, also auf die Periode, verteilt sind, und daraus ergeben sich durch eine einfache Rechnung die numerischen Werte der Koeffizienten. Je mehr dieser Koeffizienten man berechnen will, für desto mehr Werte des Argumentes ist die Berechnung der

Funktion auszuführen, in desto mehr Teile ist also der Umkreis zu teilen.

Ein ähnliches Verfahren wird bekanntlich heutzutage vielfach zur Entwicklung von beobachteten Größen benutzt, so in der Astronomie zur Entwicklung der Ausdrücke für die Teilungsfehler oder der durch die mechanische Konstruktion entstehenden Fehler eines geteilten Kreises und in der Meteorologie unter der Bezeichnung harmonische Analyse zur Darstellung irgend welcher als periodisch vorausgesetzter Erscheinungen durch eine trigonometrische Reihe sowie zu vielen anderen Zwecken. Zur besseren Übersicht mögen die Grundlagen des Verfahrens, wie es bei Gauß auftritt, hier auseinandergesetzt werden:

S sei eine periodische Funktion von x mit der Periode 2π, die sich durch folgende Reihe darstellen läßt

76) $$S = \sum \alpha_m \cos mx + \sum \beta_m \sin mx$$

und es seien die Zahlenwerte von S für eine Reihe gleichmäßig auf die Periode verteilter Werte von x, nämlich $x = 0, \omega, 2\omega \ldots (r-1)\omega$ bekannt. Diese Zahlenwerte seien $S_0, S_1, S_2 \ldots S_{r-1}$; aus ihnen können die α_m und β_m berechnet werden. Setzt man nämlich (vgl. art. 20 der Theoria interpolationis)

77) $$\begin{aligned}
\alpha_0 &= \frac{1}{r}\sum_0^{r-1} S_i \\
\alpha_1 &= \frac{2}{r}\sum_0^{r-1} S_i \cos i\omega & \beta_1 &= \frac{2}{r}\sum_0^{r-1} S_i \sin i\omega \\
\alpha_2 &= \frac{2}{r}\sum_0^{r-1} S_i \cos 2i\omega & \beta_2 &= \frac{2}{r}\sum_0^{r-1} S_i \sin 2i\omega \\
&\ldots\ldots & &\ldots\ldots \\
\alpha_m &= \frac{2}{r}\sum_0^{r-1} S_i \cos mi\omega & \beta_m &= \frac{2}{r}\sum_0^{r-1} S_i \sin mi\omega \\
&\ldots\ldots & &\ldots\ldots
\end{aligned}$$

so ist folgendes zu bemerken, wofür Gauß die Beweise erbringt:

Bricht die Reihe 76) mit einem bestimmten Gliede $m = \mu$ ab, so ist die Anzahl der zu bestimmenden Koeffizienten gleich $2\mu + 1$; ist die Zahl der gegebenen Werte von S, also r, ebenfalls gleich $2\mu + 1$, so sind die Koeffizienten nicht nur eindeutig bestimmt, sondern die Darstellung von S durch 76—77) ist auch streng richtig; ist $r = 2\mu$, also eine gerade Zahl, so sind durch die obigen Formeln alle Koeffizienten α und β, mit Ausnahme der letzten, α_μ und β_μ, streng bestimmt; die beiden letzteren bleiben unbestimmt.

Ist jedoch die Anzahl der Koeffizienten größer als die An-

zahl der gegebenen Werte S_i, ist die Reihe also im besonderen eine unendliche, so werden zwar die Koeffizienten, wenn $r = 2\mu + 1$ ist, bis α_μ, β_μ, und wenn $r = 2\mu$ ist, bis $\alpha_{\mu-1}$, $\beta_{\mu-1}$ eindeutig bestimmt; die Darstellung der Funktion S ist aber nur eine genäherte und zwar werden die Koeffizienten α, β soweit genähert gefunden, als die auf α_μ, β_μ bzw. $\alpha_{\mu-1}$, $\beta_{\mu-1}$ folgenden Koeffizienten vernachlässigt sind.

Hieraus ergibt sich für die praktische Anwendung folgender Schluß: Schreiten die Koeffizienten α, β nach Potenzen einer Größe ε fort, ist also etwa $\alpha_i = a_i\,\varepsilon^i$, $\beta_i = b_i\,\varepsilon^i$ und ist die Reihe unendlich, so gilt die Darstellung 76) genähert bis auf Größen der Ordnung ε^μ einschließlich, wenn $r = 2\mu + 1$, und ausschließlich, wenn $r = 2\mu$. Die Formeln 77) sind Werke VII, 1906 S. 496 f. für die Werte $\mu = 48$, 24, 12 für die praktische Rechnung in vereinfachte Form gebracht und durch abgedruckte Beispiele aus dem Nachlaß erläutert worden.

Gauß benutzt auch diese Methode, um die Richtigkeit seiner Konvergenzuntersuchung der Entwicklung der Mittelpunktsgleichung zu prüfen, Werke X 1, S. 424 (vgl. auch ebend. S. 432—433 u. 442).

Bei den Störungsrechnungen handelt es sich um die Entwicklung einer Funktion nach zwei Veränderlichen, den mittleren Längen des störenden und des gestörten Körpers; auf diesen Fall kommt Gauß in der Theoria interpolationis nicht zu sprechen und es finden sich darüber auch nirgends Aufzeichnungen im Nachlaß, außer den Zahlenrechnungen für die Störungen der Ceres und der Pallas. Es bietet indessen keine Schwierigkeiten, das Verfahren auch hierfür zu entwickeln:

Sei S eine periodische Funktion mit der Periode 2π von zwei Variabeln x und y, welche sich in der Form:

78) $$S = \sum\sum \alpha_{m\cdot n} \cos(mx + ny) + \sum\sum \beta_{m\cdot n} \sin(mx + ny)$$

darstellen läßt.

Seien ferner die Zahlenwerte von S bekannt für alle Wertepaare x, y, die man erhält, wenn die Periode für x in r Teile und die für y in s Teile geteilt wird; das Intervall der Teilung für x sei $\omega = \frac{2\pi}{r}$, das für y sei $\omega' = \frac{2\pi}{s}$ und die zahlenmäßig bekannten Werte von S, deren Anzahl rs ist, seien

$S_{0\cdot 0}$ für $x = 0$,	$y = 0$		$S_{1\cdot 0}$ für $x = \omega$,	$y = 0$
$S_{0\cdot 1}$ „ „	$y = \omega'$		$S_{1\cdot 1}$ „ „	$y = \omega'$
$S_{0\cdot 2}$ „ „	$y = 2\omega'$		$S_{1\cdot 2}$ „ „	$y = 2\omega'$
. . . . .	. . .		. . . . .	. . .
$S_{0\cdot s-1}$ „ „	$y = (s-1)\omega'$		$S_{1\cdot s-1}$ „ „	$y = (s-1)\omega'$

u. s. w.

Ferner seien

$S_0, S_1 \ldots S_{r-1}$ die Werte von S für $x = 0, \omega, 2\omega \ldots (r-1)\omega$;

dann sind diese Werte periodische Funktionen von y und man kann die Entwicklungen ansetzen:

79)
$$\begin{aligned}
S_0 &= p_{0\cdot 0} + p_{0\cdot 1}\cos y + p_{0\cdot 2}\cos 2y + \cdots \\
&\quad + q_{0\cdot 1}\sin y + q_{0\cdot 2}\sin 2y + \cdots \\
S_1 &= p_{1\cdot 0} + p_{1\cdot 1}\cos y + p_{1\cdot 2}\cos 2y + \cdots \\
&\quad + q_{1\cdot 1}\sin y + q_{1\cdot 2}\sin 2y + \cdots \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
S_{r-1} &= p_{r-1\cdot 0} + p_{r-1\cdot 1}\cos y + p_{r-1\cdot 2}\cos 2y + \cdots \\
&\quad + q_{r-1\cdot 1}\sin y + q_{r-1\cdot 2}\sin 2y + \cdots
\end{aligned}$$

und nach der im vorhergehenden beschriebenen Interpolationsmethode die Koeffizienten p und q in diesen Reihen zahlenmäßig berechnen und zwar

$$\begin{aligned}
&p_{0\cdot 0}, p_{0\cdot 1} \ldots;\ q_{0\cdot 1}, q_{0-2} \ldots \text{ aus } S_{0\cdot 0}, S_{0\cdot 1} \ldots S_{0\cdot s-1}; \\
&p_{1\cdot 0}, p_{1\cdot 1} \ldots;\ q_{1\cdot 1}, q_{1\cdot 2} \ldots \text{ aus } S_{1\cdot 0}, S_{1\cdot 1} \ldots S_{r-1\cdot 1} \text{ u. s. w.}
\end{aligned}$$

Setzt man weiter

80)
$$\begin{aligned}
S &= p_0 + p_1\cos y + p_2\cos 2y + \cdots \\
&\quad + q_1\sin y + q_2\sin 2y + \cdots,
\end{aligned}$$

so sind die Koeffizienten p und q in diesem Ausdruck periodische Funktionen von x und zwar sind $p_{0\cdot 0}, p_{1\cdot 0} \ldots p_{r-1\cdot 0}$ die Zahlenwerte von p_0 für $x = 0, \omega, \ldots (r-1)\omega$; $p_{1\cdot 0}, p_{1\cdot 1}, \ldots p_{r-1\cdot 1}$ die Zahlenwerte von p_1 für die gleichen x-Werte u. s. w. Es lassen sich daher in den Reihen

81)
$$\begin{aligned}
p_0 &= \gamma_{0\cdot 0} + \gamma_{0\cdot 1}\cos x + \gamma_{0\cdot 2}\cos 2x + \cdots \\
&\quad + \delta_{0\cdot 1}\sin x + \delta_{0\cdot 2}\sin 2x + \cdots \\
p_1 &= \gamma_{1\cdot 0} + \gamma_{1\cdot 1}\cos x + \gamma_{1\cdot 2}\cos 2x + \cdots \\
&\quad + \delta_{1\cdot 1}\sin x + \delta_{1\cdot 2}\sin 2x + \cdots \\
&\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

die Koeffizienten γ und δ wieder zahlenmäßig berechnen.

Es bleibt jetzt nur noch übrig, die gefundenen Reihen 81) für $p_0, p_1 \ldots$ in 80) einzusetzen und die Produkte der Cosinus und Sinus auszumultiplizieren, damit man die Entwicklung in der Form 78) erhält.

Bei der Besprechung der Pallasstörungen werden wir noch weitere Kunstgriffe kennen lernen, die Gauß bei diesen seinen Entwicklungen benutzt hat (s. unten).

Da die Entwicklung der in der Störungsfunktion auftretenden Funktionen in trigonometrische Reihen nach den Vielfachen der mittleren Längen oder Anomalien gleichzeitig nach Potenzen der Exzentrizitäten (und Neigungen) fortschreitet, ohne aber eine eigentliche Potenzreihe zu sein, so bedeutet die Einteilung der Periode in eine bestimmte Anzahl von Teilen auch die Vernachlässigung einer bestimmten Potenz der Exzentrizitäten.

Nach einer Äußerung von Gauß in einem Brief an Olbers vom 25. März 1805 plante er die Entwicklung bis zu den 11. Potenzen der Exzentrizitäten, beschränkte sich dann aber nach einem weiteren Briefe vom 10. Mai desselben Jahres auf die 5. Potenzen.

Er begann seine Rechnungen mit den Breitenstörungen der Ceres; diese sind Werke VII, 1906, S. 401 f. abgedruckt. Die hier angewandte Methode zeigt aber bei weitem noch nicht die Vollkommenheit, wie die bei den Störungen der Pallas angewandte. Ein Grund hierfür mag darin liegen, daß Gauß zunächst Koordinatenstörungen rechnete, bei denen die Störungen nicht, wie bei den Elementenstörungen, durch einfache Quadraturen gefunden werden.

Die Gleichung für die Breitenstörungen ist dieselbe wie oben, nämlich:

82) $$\Theta = \frac{\mu a}{\sqrt{1-e^2}}\left\{\cos V\int r\left(\frac{dR}{dz}\right)\sin V . n\,dt - \sin V\int r\left(\frac{dR}{dz}\right)\cos V . n\,dt\right\}$$

(vgl. auch Werke VII. 1906, S. 402).

Hier läßt sich die rechte Seite nicht direkt nach der Theoria interpolationis entwickeln, sondern es müssen erst die Größen unter den Integralzeichen entwickelt, dann die Integration ausgeführt und dann die Multiplikation mit den ebenfalls zu ent Größen $\cos V$ und $\sin V$ ausgeführt werden.

Die Größen unter dem Integralzeichen werden durch trigonometrische Reihen dargestellt, die nach zwei Argumenten, den mittleren Längen des störenden und des gestörten Körpers, fortschreiten.

Um das etwas verwickelte von Gauß hier angewandte Verfahren klar zu machen, wollen wir das Glied unter dem ersten Integralzeichen, nämlich

$$X = r\left(\frac{dR}{dz}\right)\sin V$$

betrachten. Es ist, wie oben Gl. 53) (Werke VII, 1906, S. 402)

$$\left(\frac{dR}{dz}\right) = z'\left(\frac{1}{r'^3} - \frac{1}{\mathfrak{R}^3}\right)$$

und [1])

$$z' = r' \sin J \sin(v' - \Omega) = r' \sin J \sin(♃ + \varepsilon' - \Omega).$$

Hiermit wird:

$$X = X_1 X_2$$

83) $$X_1 = rr' \sin J . \sin V . \sin(♃ + \varepsilon' - \Omega)$$

$$X_2 = \left(\frac{1}{r'^3} - \frac{1}{\mathfrak{R}^3}\right).$$

Die Größe $\sin J$ wird als konstant angesehen, da die vom Quadrat der störenden Masse abhängenden Glieder nicht berücksichtigt werden; die Größen r und $\sin V$ können nach Vielfachen der mittleren Anomalie M entwickelt werden in der Form

84) $$\sum a_n \begin{matrix}\cos\\ \sin\end{matrix} nM,$$

wo die a_n von der Form

85) $$a_n = e^n \sum b_i e^i$$

sind. Die Reihe nach den Vielfachen von M schreitet also nach Potenzen von e fort, aber in der Weise, daß a_n nicht gleich dem Produkt aus e^n in einen konstanten Faktor, sondern selbst gleich einer Potenzreihe nach e ist, welche mit dem Gliede in e^n beginnt. Da die analytische Entwicklung einerseits zu ziemlich unübersichtlichen Ausdrücken führt (man kann sie am besten durch Anwendung der Besselschen Funktionen ausführen), andererseits bei nicht allzu großem e die Glieder hinreichend stark abnehmen, so empfiehlt sich hier die interpolatorische Methode. Das gleiche gilt von der Entwicklung der Größe r' und von der der Größe $\sin(♃ + \varepsilon' - \Omega)$, die nach Vielfachen der mittleren Anomalie Jupiters M' und gleichzeitig nach Potenzen der Exzentrizität e' fortschreiten.

Gauß entwickelt daher im Prinzip X_1 nach der interpolatorischen Methode und teilt den Umkreis in 10 Teile, wobei er die Rechnung aber für die Werte M und M' gleich 18°, 54°, 90°, ... 342° ausführt. Hierdurch vereinfacht sie sich. Die Formeln 77) bedürfen dazu einer Umformung, die Bd. VII, 1906, S. 404 ange-

---

1) J Neigung beider Bahnen gegeneinander
v' wahre Länge Jupiters
Ω Knotenlänge
ε' Mittelpunktsgleichung Jupiters.

geben ist. Die Einteilung des Umkreises in 10 Teile entspricht nach obigem der Vernachlässigung der 5. Potenzen der Exzentrizitäten.

Anders liegt die Sache bei der Größe $\frac{1}{\mathfrak{R}^3}$; es ist

$$\mathfrak{R} = (r^2 + r'^2 - 2rr'\cos w)^{\frac{1}{2}}.$$

Setzt man der Kürze halber[1]) $x =$ ♃ $-$ ⚳ $+ \varepsilon' - \alpha$, so ist $\cos w = \cos b \cos x$ und man kann setzen (Werke VII, 1906, S. 404)

86) $$\mathfrak{R}^2 = \frac{rr'\cos b}{\beta}(1 + \beta^2 - 2\beta\cos x),$$

wo

87) $$\beta = \frac{\sqrt{r^2 + r'^2 + 2rr'\cos b} - \sqrt{r^2 + r'^2 - 2rr'\cos b}}{\sqrt{r^2 + r'^2 + 2rr'\cos b} + \sqrt{r^2 + r'^2 - 2rr'\cos b}}$$

Gauß entwickelt

88) $$\frac{2}{\mathfrak{R}^3} = \frac{1}{2}Q + Q'\cos x + Q''\cos 2x + \cdots$$

und erhält hiermit

89) $$-2X_2 = \frac{2}{\mathfrak{R}^3} - \frac{2}{r'^3} = \frac{1}{2}Q^{(0)} + Q'\cos x + \cdots,$$

wo $Q^{(0)} = Q - \frac{4}{r'^3}$ und wo die Q außer von dem Faktor $\frac{rr'\cos b}{\beta}$ nur noch von β abhängen. Da $\cos b$ nahe gleich Eins ist, so ist β nahe gleich $\frac{r}{r'}$, da wir $r' > r$ voraussetzen. Ist nun das Verhältnis $\frac{r}{r'}$ nicht erheblich von Eins verschieden, so nehmen die Glieder der Reihe weniger stark ab; denn sie schreiten im wesentlichen nach Potenzen von β fort. Es ist wünschenswert bei der Entwicklung ziemlich weit zu gehen. Hierin wird der Grund zu suchen sein, warum Gauß den Faktor X_2 analytisch entwickelt. Wahrscheinlich hat er schon hier die im Briefe an Bessel (Werke X 1, S. 238, vergl. hier unten S. 81) beschriebene Methode zur Berechnung der Q, wenigstens teilweise, benutzt.

Die Ausführung der ganzen etwas unübersichtlichen Rechnung gestaltet sich etwa so:

Ist zunächst die Entwicklung 89) von X_2 nach der analytischen

---

1) α Mittelpunktsgleichung der Ceres minus Reduktion auf die Jupitersbahn. b heliozentrische Breite der Ceres.

Methode für jede Wertekombination r, r', $\cos b$, also für jedes Wertepaar M, M', das für die interpolatorische Methode gebraucht wird, durchgeführt, so sind damit die Zahlenwerte der Q für jedes dieser Wertepaare, 100 an der Zahl, bekannt (vgl. die Werte von $\log Q''$ in der Tabelle Werke VII, 1906, S. 405). Bei der ersten Methode der Ceresstörungen (oben Gl. 55 u. f.) war diese Entwicklung nur einmal für die konstanten Größen a und a' auszuführen. Gauß hat hier die Q bis zum Gliede $Q^{(10)}$ berechnet, also dabei noch die 10te Potenz von $\frac{r}{r'}$ berücksichtigt; denn es ist $r'^3 Q^{(n)}$ von der Größenordnung $\left(\frac{r}{r'}\right)^n$.

Setzt man die Reihe 89) für X_2 in den Ausdruck 83) für X ein, so erhält man z. B. für den mit Q'' multiplizierten Teil von X (vgl. auch VII, 1906, S. 403, Gleichung 2):

$$90)\qquad \text{pars } X = \frac{rr'}{4} \sin J \sin V . Q'' \{ \sin(D - \text{⚳} + \varepsilon' - 2\alpha + \Omega) - \sin(3D + \text{⚳} + 3\varepsilon' - 2\alpha - \Omega) \} .$$

Die Zahlenwerte dieser Glieder werden wieder für jedes Wertepaar M, M' berechnet und damit durch die Interpolation nach den Formeln 77)—81) die Entwicklung dieses Teils von X gefunden.

Die Einzelheiten der Gaußschen Rechnungsart, die hier nur in großen Umrissen geschildert ist, sind in Werke VII, 1906, S. 403—406 an dem zweiten Gliede des vorstehenden Ausdrucks 90) als Beispiel auseinandergesetzt.

Ist die Größe X nach Vielfachen von M und M' entwickelt, so ist nach 82) die Integration durch Anbringung der Divisoren auszuführen und sodann mit der gleichfalls auf interpolatorischem Wege gewonnenen Entwicklung von $\cos V$ und dem konstanten Faktor $\frac{\mu a}{\sqrt{1-e^2}}$ zu multiplizieren.

Wie bereits erwähnt, hat Gauß nach der eben besprochenen Methode nur die Breitenstörungen gerechnet, die er in einem Briefe an Olbers vom 2. Juli 1805 mitteilt (vgl. auch Bd. VII, 1906 S. 408—409). Diese Störungen enthalten im ganzen 27, nach der Zusammenziehung der Glieder gleicher Periode 16 Glieder.

Die Rechnung wird vereinfacht durch verschiedene auftretende Symmetrien, insbesondere auch durch die der beiden Glieder in Gleichung 82), von denen wir oben nur das erste betrachtet haben. Sie ist aber dennoch äußerst umständlich. Die Berechnung der Breitenstörungen dehnt sich trotz der raumersparenden Schreib-

weise von Gauß, bei der durchschnittlich 2000—3000 Ziffern auf einer Quartseite des Handbuchs Bc stehen, über 11 Seiten aus.

Aus seinen Untersuchungen über Ceres schließt Gauß im genannten Briefe vom 2. Juli 1805, daß „die sämtlichen bei der Pallas merklichen Gleichungen der Länge, Breite und des Radiusvektor leicht auf 500 steigen können“ und „daß man vielleicht mehrere Tage nötig hat, um einen einzigen Ort zu berechnen“ und fährt fort: „Durch ähnliche, aber freilich viel zahlreichere Tafeln, wie meine älteren[1]) für die Ceresstörungen, wird indeß die Arbeit sehr erleichtert werden können, aber die Berechnung solcher Tafeln wird auch Monate kosten, nachdem die Formeln entwickelt sind“.

VI. Tafeln zur Entwicklung von $(a^2 + a'^2 - 2aa' \cos x)^{-\frac{1}{2}}$.

Gauß hatte wohl zunächst die Absicht, nach der im vorigen geschilderten Methode oder nach einer ganz ähnlichen die Störungen der Ceres zu vollenden und auch die der anderen Planeten (der Pallas und der im März 1805 entdeckten Juno) zu berechnen. Die Berechnung der Breitenstörungen der Ceres hatte gute Ergebnisse geliefert, war aber recht umfangreich. Gauß suchte daher die Arbeit durch Entwerfung von Tafeln abzukürzen und es lag am nächsten, solche Tafeln für die Entwicklung des Ausdruckes 88), also für die Koeffizienten Q herzustellen.

Bessel hatte ihm seine Hilfe bei der Ausführung von Zahlenrechnungen angeboten und auch schon bei einer anderen Gelegenheit gewährt.

Nach den obigen Gleichungen 83) und 86) galt es, den Ausdruck

$$(1 + \beta^2 - 2\beta \cos x)^{-\frac{3}{2}}$$

für verschiedene Werte von β zu entwickeln. Waren dann die Koeffizienten Q für eine Reihe von solchen Werten tabuliert, so konnten diese Tafeln auch für andere Planeten angewandt werden. In dem Briefe an Bessel vom 3. Sept. 1805 (Werke X 1, S. 237) bittet Gauß diesen, ihn bei der Berechnung der Tafel zu unterstützen und setzt ihm die Berechnungsweise eingehend auseinander.

Die Vorschriften, die er Bessel gibt, beziehen sich allerdings auf die Entwicklung von

$$(1 + \beta^2 - 2\beta \cos x)^{-\frac{1}{2}}$$

1) Vgl. oben S. 72.

bezw.

$$(a^2 + a'^2 - 2aa' \cos x)^{-\frac{1}{2}} = \tfrac{1}{2} A^{(0)} + A' \cos x + A'' \cos 2x + \cdots.$$

Es ist wohl anzunehmen, daß Gauß diesen Ausdruck gewählt hat, weil er ihn auch sonst brauchte und die Entwicklung, die zu den Exponenten $-\frac{3}{2}$, $-\frac{5}{2}$ usw. gehört, sich leicht aus der zu $-\frac{1}{2}$ gehörenden herstellen läßt.

Die Entwicklungsmethode, die Gauß im Brief an Bessel schildert, ist bei weitem vollkommner als die bei der ersten Berechnung der Ceresstörungen (S. 67 f.) angewandte. Es ist wohl anzunehmen, daß Gauß gerade im Anschluß an diese Untersuchungen begonnen hat, sich mit der Theorie der allgemeinen hypergeometrischen Reihe $F(\alpha, \beta, \gamma, x)$ zu beschäftigen; die ältesten uns erhaltenen Aufzeichnungen zu dieser Theorie stammen aus den ersten Monaten des Jahres 1806, sie sind Werke X 1, S. 326 abgedruckt (vergl. die Bemerkungen von Schlesinger, a. a. O., S. 331). Man findet in diesen Aufzeichnungen den Kettenbruch, der den Quotienten zweier hypergeometrischen Reihen darstellt (siehe a. a. O., S. 329, Verwandlung), und der den in dem Briefe an Bessel gegebenen Kettenbruch für den Quotienten zweier aufeinanderfolgender Q als besonderen Fall enthält. Den hierauf gegründeten Beweis für die an Bessel mitgeteilten Entwicklungen gibt Gauß in einer im Jahre 1809 verfaßten Aufzeichnung „Einiges über die unendliche Reihe usw.“, die Werke X 1, S. 338 abgedruckt ist (siehe insbesondere a. a. O., S. 345—349 und die zugehörigen Bemerkungen von Schlesinger, S. 354). Im Nachlaß und in Veröffentlichungen finden sich noch viele Stellen, auch aus späterer Zeit, die von der Entwicklung des obigen Ausdrucks handeln; so in den Disquisitiones circa seriem infinitam etc. (1812) Werke III, S. 128, ferner Werke VIII, S. 84; vergl. auch den Aufsatz von Schlesinger über Gauß' Arbeiten zur Funktionentheorie, S. 84, 90—91. Die fertig gerechnete Tafel befindet sich im Nachlaß; sie gibt die Größen $\pi^{(i)} = \left(\frac{a}{a'}\right)^i \sqrt{a'^2 - a^2}\, A^{(i)}$ mit dem Argument $\Theta = \operatorname{arc\,tang} \frac{a}{a'}$ und geht bis $\pi^{(10)}$. Sie scheint aber niemals zur Anwendung gelangt zu sein, da Gauß, als er nach langer Unterbrechung die Störungen der Pallas im Jahre 1810 vornahm, das rein interpolatorische Verfahren anwandte.

VII. Störungen der Pallas und Vorbereitungen dazu.

1. Disquisitio de elementis ellipticis Pålladis.

Diese Abhandlung ist im wesentlichen als eine Vorarbeit zu den Störungsrechnungen anzusehen.

Die zur Anwendung kommenden Methoden sind bereits oben S. 41 f. besprochen; nach ihnen bestimmt Gauß zunächst rein elliptische Elemente der Pallas aus den drei folgenden Gruppen von Oppositionen: 1803, 04, 05, 07; 1804, 05, 07, 08; 1805, 07, 08, 09.

Es zeigt sich einerseits aus der Vergleichung mit den Beobachtungen, daß für den Zeitraum von vier Oppositionen rein elliptische Elemente die Bahn genau genug darstellen, um den Planeten für die nächste Folgezeit so vorauszuberechnen, daß eine Wiederauffindung mit dem Fernrohr keine Schwierigkeit macht; andererseits aber weichen die drei Elementensysteme doch so von einander ab, daß die Berücksichtigung der Störungen als eine Forderung der Wissenschaft bezeichnet werden muß (Disqu. art. 9—10).

Es handelte sich nun um die Frage, welches Elementensystem der Störungsrechnung zugrunde gelegt werden sollte, und Gauß wählt dasjenige (Disqu. art. 11 f.), das sich allen beobachteten Oppositionen am besten anschließt. Von der Bestimmung eines solchen handeln bereits die art. 172—189 der Theoria motus, in denen Gauß die Methode der kleinsten Quadrate ausführlich begründet.

Die hier benutzten Gleichungen zur Bestimmung der Verbesserungen der Elemente erhalten, wenigstens mit bezug auf die Längen, eine besondere Form, weil es sich um Oppositionen handelt, also die heliozentrische Länge statt der geozentrischen gebraucht wird; sie sind oben S. 47 abgedruckt.

Am Schluß der Disqu. gibt Gauß an, wie das gefundene Elementensystem die sechs Normalorte darstellt; es zeigen sich in der Länge Differenzen bis zu $3^1/_2$ Minute und in der Breite bis fast $1^1/_2$ Minute.

2. Spezielle Störungen der Pallas.

Die Rechnung der speziellen Störungen betrachtet Gauß als Vorbereitung zur Berechnung der allgemeinen Störungen; die von ihm angewandte Methode ist heute wenigstens in ihren Grundzügen jedem theoretischen Astronomen geläufig.

Gauß berechnet die Störungen der Elemente: Neigung, Knotenlänge, mittlere tägliche Bewegung, Exzentrizitätswinkel, Perihel-

länge und mittlere Länge in der Epoche; die Differentialgleichungen (Werke VII, 1906, S. 474) unterscheiden sich nicht wesentlich von den in der nachgelassenen Abhandlung „Exposition d'une nouvelle méthode de calculer les perturbations planétaires" (Werke VII, S. 457) gegebenen.

Die von ihm gebrauchten Formeln für die mechanische Quadratur leitet Gauß ebenfalls in der genannten „Exposition" (Werke VII, 1906, S. 462) ab. Weitere Nachlaßnotizen darüber sind nicht vorhanden und auch die numerischen Rechnungen sind nur ganz lückenhaft erhalten. Die mechanische Quadratur hängt bekanntlich mit der Interpolationsrechnung zusammen. Über die erstere hat Gauß eine Abhandlung „Methodus nova integralium valores usw." 1814 veröffentlicht, die aber in keiner näheren Beziehung zu seinen Störungsrechnungen steht; seine Arbeiten über die letztere sind erst aus der von Schering in Werke III aus dem Nachlaß abgedruckten Abhandlung „Theoria interpolationis methodo nova tractata" bekannt geworden; in dieser geht Gauß nicht auf mechanische Quadratur ein, sondern entwickelt nur die Grundlage für seine Rechnung der allgemeinen Störungen. Dagegen hat Encke nach Vorlesungen von Gauß im Berliner astronom. Jahrbuch für 1830 in einem Aufsatz „Über Interpolation" [1]) einiges gegeben, das in anderer Form in der „Theoria interpolationis" zu finden ist, und in einem zweiten Aufsatz „Über mechanische Quadratur" im Jahrbuch für 1837 die Methode auseinandergesetzt, die ihm Gauß bei Übertragung der Berechnung der speziellen Störungen der Pallas im Jahre 1812 mitgeteilt habe [2]). Hiernach scheint es, daß

1) Encke leitet den Aufsatz mit der Bemerkung ein: Der folgende Aufsatz ist aus den Vorlesungen entlehnt, die ich im Jahre 1812 bei dem Herrn Hofrath Gauß zu hören das Glück hatte. In dem ganzen Gange der Entwicklung bin ich, so viel die Erinnerung gestattete, dem Vortrage meines geehrten Herrn Lehrers gefolgt, da er die größte Gründlichkeit mit der größten Einfachheit und Eleganz verbindet.

2) Encke bemerkt am Anfange dieses Aufsatzes: Bei meinem Aufenthalt in Göttingen im Jahre 1812 übertrug mir Herr Hofrath Gauß die Berechnung der speziellen Störungen der Pallas, und leitete mir zu diesem Behufe seine Methoden und Formeln ab, deren er seit längerer Zeit sich bedient hatte. Er hatte damals die Absicht selbst etwas über diesen Gegenstand bekannt zu machen und behielt sich diese Erläuterung vor. Jetzt wo leider die Aussicht auf ein eigenes Werk von Gauß, wegen seiner vielfachen andern wichtigen Untersuchungen, so gut wie verschwunden scheint, hat er es mir gestattet, das was ich aus seinen Vorträgen für die nachherige häufige Anwendung auf Kometen und kleine Planeten benutzt habe hier zu publiziren; wobei ich nur noch hinzuzufügen mir erlaube, daß der Weg zum Beweise der Formeln nicht genau der ist, welchen Gauß bei mir genommen, weil es mir nicht ratsam schien, allzuviele verwandte Betrachtungen ein-

Encke die Fortsetzung der Gaußschen Rechnung der speziellen Störungen von 1811 übernommen hat; denn die in Werke VII, 1906 abgedruckten bis 1811 reichenden Störungsrechnungen sind im Jahre 1810 von Gauß selbst ausgeführt. Von der Enckeschen Berechnung ist anscheinend nichts erhalten.

Die erste Rechnung der speziellen Störungen der Pallas beruhte während des ganzen sich von Juni 1803 bis September 1811 erstreckenden Zeitraums auf einem konstanten Elementensystem, nämlich dem in der „Disquisitio de elementis Palladis“ abgeleiteten; daher sind dabei nur die Störungen von der ersten Ordnung berücksichtigt; die Verbesserungen der Elemente sind nach der oben S. 45—46 besprochenen Methode ermittelt.

Um noch schärfer zu rechnen und die Störungen zweiter Ordnung zu berücksichtigen, wiederholt Gauß die Rechnung mit veränderlichen Elementen, indem er „den ganzen Zeitraum 1803—1811 in acht Perioden“ teilt (vgl. Brief an Olbers vom 13. Dez. 1810 und Bd. VII, 1906, S. 464) und jeder besondere aus der ersten Rechnung abgeleitete Elemente zugrunde legt (Bd. VII, S. 483). Weiteres siehe oben S. 57.

3. Erste Rechnung der allgemeinen Störungen der Pallas.

Hier berechnet Gauß, wie bei den speziellen Störungen, die Störungen der Elemente. Die Differentialgleichungen (Werke VII, 1906, S. 490) zeigen einige nicht sehr wesentliche Unterschiede gegen die früher bei Ermittlung der speziellen Störungen gebrauchten; doch führt Gauß anstelle des Exzentrizitätswinkels den Logarithmus des Parameters und anstelle der mittleren täglichen Bewegung ihren hyperbolischen Logarithmus bezw. den gemeinen Logarithmus der halben großen Axe ein.

Die Ausdrücke für die Ableitungen der Elemente, die, nachdem sie durch trigonometrische Reihen dargestellt sind, direkt integriert werden können, sind periodische Funktionen der beiden Argumente M und M', d. i. der mittleren Anomalien von Pallas und Jupiter, wenn man ein konstantes elliptisches Elementensystem zugrunde legt, also die Störungen zweiter Ordnung vernachlässigt; im strengen Sinne treten noch die Apsiden- und Knotenbewegungen als Argumente hinzu.

---

zumischen. Diese Bemerkung soll, wie sich von selbst versteht, nur bevorworten daß wenn vielleicht in der Beweisführung Einiges nicht bestimmt genug erscheinen möchte, der Fehler ganz allein mir Schuld gegeben werden muß.

Gauß beschränkt sich durchweg auf die Ermittlung der Störungen erster Ordnung; dagegen bemüht er sich, möglichst hohe Potenzen der Exzentrizitäten und des Verhältnisses der mittleren Entfernungen zu berücksichtigen. Er wandte daher die oben S. 75—76 besprochene rein interpolatorische Methode an. Vielleicht hat er auch die in den artt. 25—28 der Theoria interpolationis entwickelte Methode benutzt, bei der die Periode in eine Anzahl primärer Teile und diese wieder weiter sekundär eingeteilt wird, wodurch das Schema der Rechnung vereinfacht wird.

Da jedoch im Nachlaß nur die numerischen Rechnungen für Pallas ohne jedes erklärende Wort vorhanden sind, so wurden in Werke VII, 1906, S. 491—500 alle Formeln hergestellt, die Gauß nach Ausweis seiner numerischen Rechnungen benutzt hat.

Über die Anordnung der Gaußschen Rechnung und über seine Kunstgriffe ist folgendes zu sagen:

Die Ausdrücke der Ableitungen der Elemente sind zwar periodische Funktionen der beiden Argumente M und M' und können nach den Formeln auf S. 75—76 entwickelt werden; ihre gesamte Entwicklung nach diesen Argumenten setzt sich aber aus drei Einzelentwicklungen zusammen: die Koordinaten der Pallas sind periodische Funktionen von M und die des Jupiter ebensolche von M', die gegenseitige Entfernung beider Planeten aber ist eine periodische Funktion von $M' - M$. Die Glieder in den Entwicklungen der Pallaskoordinaten nehmen ab wie die Potenzen der Pallasexzentrizität und die der Jupiterkoordinaten wie die der Jupiterexzentrizität; da die erstere groß und die letztere klein ist, so folgt zunächst, daß man bei der letzteren nicht so weit in der Entwicklung zu gehen braucht, wie bei der ersteren, also für das Argument M den Umkreis in eine größere Anzahl von Abschnitten teilen wird, als für M'; die Entwicklung nach diesem letztern Argument tritt überhaupt in den Hintergrund.

Die Entwicklung des gegenseitigen Abstandes beider Planeten nach Vielfachen von $M' - M$ enthält ebenfalls die Potenzen beider Exzentrizitäten, aber außerdem nehmen ihre Glieder nach Potenzen der Größe α ab, wo $\alpha = \frac{a}{a'}$ das Verhältnis der beiden halben großen Axen ist, und diese letztere Entwicklung ist also ebenfalls viel weiter fortzusetzen, als die nach M'. Mit ihr hat sich Gauß schon bei seinen früheren Untersuchungen im Anschluß an seine Arbeiten über das arithm.-geom. Mittel besonders beschäftigt, wie oben besprochen worden ist. Er hat sie bei den Ceresstörungen im wesentlichen noch nach der analytischen Methode ausgeführt.

Es sind also die Entwicklungen nach M und $M' - M$ die wichtigen und aus diesem Grund hat er wahrscheinlich diese beiden Argumente gewählt; denn bei dieser Anordnung gehen die größten Glieder den kleineren voraus. Gauß hat in seiner ersten Rechnung der allgemeinen Störungen der Pallas den Umkreis für M in 12 und für $M' - M$ in 48 Teile geteilt, wodurch also im großen und ganzen die 5. Potenzen der Exzentrizität und die 23. Potenzen von α berücksichtigt werden. In betreff der Einzelheiten der Rechnung, im besonderen der Zusammenziehung der Cosinus- und Sinusglieder in ein einziges Glied wird ein Hinweis auf Bd. VII genügen; die Ergebnisse für die Störungen findet man ebenfalls dort, ebenso wie die Vergleichung mit den Oppositionen und die Verbesserung der Elemente; die letztere allerdings unvollständig, da sich im Nachlaß nur weniges darüber vorfand.

4. Zweite Rechnung der allgemeinen Störungen der Pallas.

Die zweite Rechnung der allgemeinen Störungen ist noch umfangreicher als die erste, da er hier für M den Umkreis in 24 Teile teilt, also noch die 11. Potenzen von e berücksichtigt. Die Differentialgleichungen für die Elemente sind nicht wesentlich geändert. Die schließlichen Störungsausdrücke unterscheiden sich in der Form von denen der ersten Rechnung dadurch, daß Gauß hier überall in den Argumenten die mittleren Längen statt der mittleren Anomalien eingeführt und die Konstanten in den Argumenten so gewählt hat, daß alle Glieder positives Vorzeichen erhalten. Die Unterschiede der numerischen Ergebnisse in beiden Rechnungen sind klein und es wird hier ein Hinweis auf Band VII und auf S. 59—60 oben genügen.

5. Die größten Gleichungen in den Störungen der Pallas. Libration.

Ist E irgend eines der oskulierenden Bahnelemente des gestörten Planeten, so finden sich nach der Entwicklung der Störungsfunktion und der von ihr abhängenden Größen bekanntlich Gleichungen der Form:

$$91)\qquad \frac{dE}{dt} = \sum_m \sum_{m'} \alpha_{m \cdot m'} \sin(m'L' - mL + A_{m \cdot m'}),$$

wo bei Vernachlässigung der Störungen höherer Ordnung $\alpha_{m \cdot m'}$ und

und $A_{m \cdot m'}$ Konstanten sind und[1])

92)
$$L' = n't + \varepsilon',$$
$$L = \int_0^t n dt + \varepsilon$$

ist. Durch Integration findet man, wenn man n als konstant ansieht, was der Vernachlässigung der Störungen von höherer Ordnung entspricht:

93)
$$E = E_0 + \delta E,$$
$$\delta E = \alpha_{0 \cdot 0} t - \sum \frac{\alpha_{m \cdot m'}}{m'n' - mn} \cos(m'L' - mL + A_{m \cdot m'}).$$

E_0 bedeutet den mittleren Wert des Elementes und δE seine Störungen. $\alpha_{0 \cdot 0} t$ ist das sekulare Störungsglied. Die Größe der periodischen Störungsglieder hängt vom Divisor $m'n' - mn$ ab und zwar hat das größte auftretende Störungsglied im Falle der Pallas das Argument $5L' - 2L$; die fünfmalige mittlere Bewegung Jupiters minus der doppelten der Pallas $= 5n' - 2n$ ist nur gleich etwa 42″.

Gauß bezeichnet die entsprechenden Glieder unter Hinzuziehung der vom doppelten und dreifachen Argument ($10L' - 4L$ und $15L' - 6L$) abhängenden in der Länge als die große Gleichung (Werke VII, 1906, S. 559). Obwohl er bei Ermittlung der Störungen selbst nur die erste Potenz der störenden Masse berücksichtigt, berechnet er zur Bestimmung der Störungsbeträge für die einzelnen Epochen die mittlere Länge der Pallas mit Einschluß dieser großen Gleichung, nimmt also hier auf die höheren Potenzen der Störungen Rücksicht; auch zur Berechnung der mittleren Länge Jupiters benutzt er die dortige große Gleichung nach den Tafeln von Bouvard (vgl. Werke VII, 1906, S. 559—560).

Besonderes Aufsehen erregte aber die Gaußsche Mitteilung seiner Entdeckung, daß die mittleren Bewegungen von Jupiter und Pallas im rationalen Verhältnis 7 : 18 ständen, „was sich durch die Einwirkung Jupiters immer genau wieder herstellt, wie die Rotationszeit unseres Mondes". So schreibt er an Bessel[2]) am

---

1) L' = Mittlere Länge Jupiters,
L = „ „ der Pallas,
n' = konstante mittlere tägliche Bewegung Jupiters,
n = oskulierende „ „ „ der Pallas.
$\varepsilon', \varepsilon$ = Werte von L' und L für $t = 0$.

2) Werke VII, 1906, S. 421.

5. Mai 1812, als er gerade seine zweite Rechnung der allgemeinen Störungen begonnen hatte.

Aufzeichnungen darüber, wie Gauß die Entdeckung ursprünglich gemacht, sind garnicht vorhanden; wohl aber haben sich im Nachlaß nach langem Suchen zwei kleine Zettel gefunden, die aus der Zeit der Vollendung der zweiten Rechnung der allgemeinen Störungen stammen und auf denen einige numerische Rechnungen über das Librationsglied stehen, die in Bd. VII, S. 558—559 mit den nötigen Ergänzungen abgedruckt sind. Im Artikel [24] der „Exposition“ (Bd. VII, S. 468), der ebenfalls später entstanden ist, spricht Gauß von dem Fall des rationalen Verhältnisses der mittleren Bewegungen; er sagt, daß in diesem Fall in $\int n\,dt$ (also in der mittleren Länge) ein dem Quadrat der Zeit proportionales Glied entsteht, das aber periodische Form erhält, wenn man die höheren Potenzen der störenden Masse berücksichtigt und er fügt hinzu, daß dieser Fall bei Pallas wirklich eintritt. Der Inhalt der erwähnten Zettel und die Einzelausführung der zweiten Rechnung der allgemeinen Störungen geben hierfür die Erklärung. Es ist darüber folgendes zu bemerken:

Gauß kannte, als er jenen Brief an Bessel schrieb, noch nicht die Zahlenwerte der Konstanten des Librationsgliedes, auf die sich die Rechnung auf den Zetteln stützt; er fand diese vielmehr erst aus der erweiterten Entwicklung der Störungsfunktion bei der zweiten Rechnung im Hochsommer 1812 (zwischen dem 20. Juli und 10. August, vergl. Werke VII, 1906, S. 608). Man muß daher vermuten, daß er zunächst nur die Integrationsdivisoren $m'n' - mn$ zahlenmäßig untersucht hat.

Bezeichnen n_0 und ε_0 Mittelwerte von n und ε, so ist

$$L = n_0 t + \varepsilon_0 + \int \delta n\,dt + \delta\varepsilon.$$

Das Integral $\int \delta n\,dt$ enthält kein der Zeit proportionales Glied; um aber den sekularen Teil von $\delta\varepsilon$ nach den Ergebnissen der ersten Störungsrechnung festzustellen, hat man (Werke VII, 1906, S. 526, 519, 521)

sekularer Teil von $\delta\varepsilon - (1-\cos i)\,\delta\Omega - (1-\cos\varphi)\{\delta\tilde{\omega} - (1-\cos i)\,\delta\Omega\}$ $= -0''.04183$

„ „ „ $\delta\Omega$ $= -0''.09645$

„ „ „ $\delta\tilde{\omega} - (1-\cos i)\,\delta\Omega$ $= -0''.01310$

und hieraus ergibt sich

$$\text{sekularer Teil von } \delta\varepsilon = -0''.0593.$$

Da $n_0 = 769''.2443$ (Werke VII, 1906, S. 527) gefunden war, so wird der sekulare Teil von L gleich $(769''.2443 - 0''.0593)\,t$, also die augenblickliche mittlere Bewegung gleich $769''.1850$; die tropische mittlere Bewegung Jupiters hatte Gauß (Werke VII, 1906, S. 508) nach Laplace zu $299''.2650$ angenommen und mit dem ebenfalls von Gauß gebrauchten Werte der täglichen Präzession von $0''.13717$ folgt hieraus die siderische mittlere Bewegung Jupiters zu $299''.1278$. Um die kleinsten Divisoren aufzufinden, kann man das Verhältnis der mittleren Bewegungen $\frac{n'}{n} = \frac{299''.1278}{769''.1850}$ in einen Kettenbruch verwandeln und erhält diesen gleich

$$\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{3+\cfrac{1}{7913+\cdots}}}}}$$

Die Näherungswerte sind

$$\frac{1}{2};\quad \frac{1}{3};\quad \frac{2}{5};\quad \frac{7}{18}\quad \text{usw.}$$

Hier tritt außer dem Näherungsbruch $\frac{2}{5}$, der die sogenannte große Gleichung veranlaßt, ganz besonders der nächste im Betrage von $\frac{7}{18}$ hervor; dieser stellt den Bruch fast genau dar, da das darauf folgende Glied des Kettenbruchs ungewöhnlich groß ist; es ist in der Tat:

$$\frac{18}{7}\text{-mal Bewegung Jupiter} = 769''.1857$$
$$\text{Bewegung der Pallas} = 769''.1850.$$

Der Divisor $18n' - 7n$ ergab sich also verschwindend klein, und die analytische Untersuchung dieses Falls dürfte Gauß in ähnlicher Weise gemacht haben, wie Laplace bei Betrachtung der Libration der Jupitermonde.

Nach Werke VII, 1906, S. 542 ist mit den dortigen Bezeichnungen

$$\frac{dn}{dt} = 3n^2 aeT \sin(v - \pi) + \frac{3n^2 b^2}{r^2} V.$$

Die Entwicklung des Ausdruckes rechter Hand in der Form 91) möge das kritische Glied

94) $$n^2 \alpha \sin(m'L' - mL + A_{m.m'})$$

enthalten, wo in unserem Falle $m' = 18$, $m = 7$. Setzt man

95) $$u = m'L' - mL + A_{m \cdot m'}$$

und bedenkt man, daß

$$\frac{dL}{dt} = n + \frac{d\varepsilon}{dt}, \qquad \frac{dL'}{dt} = n',$$

so wird

96) $$\frac{du}{dt} = m'n' - mn - m\frac{d\varepsilon}{dt}$$

und

97) $$\frac{d^2u}{dt^2} = -mn^2\alpha \sin u + P,$$

wo P nur weniger wichtige periodische Glieder enthält. Vernachlässigt man P, und nimmt man den Faktor n^2 als konstant an, so kann die letztere Gleichung genau wie die Pendelgleichung diskutiert werden und zwar bei Annahme einer anziehenden oder abstoßenden Schwerkraft, je nachdem α positiv oder negativ ist; man erhält

98) $$\left(\frac{du}{n\,dt}\right)^2 = C + 2m\alpha \cos u.$$

Die einzelnen Fälle der Bewegung von u mögen hier aufgeführt werden, da eine wirklich vollständige Darstellung in der Literatur kaum aufzufinden sein dürfte. Ist

1. $C < -2m|\alpha|$,

so findet keine reelle Bewegung statt. Ist

2. $-2m|\alpha| < C < 0$,

so schwankt u periodisch mit der Amplitude $\pm \arccos \frac{C}{2m|\alpha|}$ und zwar um die stabile Ruhelage $u = 0$ oder π, je nachdem α positiv oder negativ ist; die Amplitude ist kleiner als $\frac{\pi}{2}$. Ist

3. $0 < C < 2m|\alpha|$,

so schwankt u in gleicher Weise mit der Amplitude

$$\pm \arccos\left(-\frac{C}{2m|\alpha|}\right),$$

die größer als $\frac{\pi}{2}$ ist. Ist endlich

4. $C > 2m|\alpha|$,

so kann $\frac{du}{dt}$ niemals verschwinden; es tritt der Fall des ganz

herumschwingenden Pendels ein, und zwar in der einen oder anderen Richtung, je nach dem Vorzeichen des Anfangswertes von $\frac{du}{dt}$.

Daneben gelten die Spezialfälle

$$\text{a.} \quad C = -2m|\alpha|;$$

stabile Ruhelage bei $u = 0$ oder π, je nachdem α positiv oder negativ.

$$\text{b.} \quad C = 0;$$

periodische Schwankung von u mit der Amplitude $\pm\frac{\pi}{2}$.

$$\text{c.} \quad C = 2m|\alpha|;$$

in diesem Falle ist

bei positivem α: $\operatorname{tg}\left(45^0 + \frac{u}{4}\right) = e^{\pm\sqrt{m\alpha}\,.\,n(t-t_0)}$,

bei negativem α: $\operatorname{tg}\frac{u}{4} = e^{\pm\sqrt{m|\alpha|}\,.\,n(t-t_0)}$.

Die Bewegung ist hier doppelt asymptotisch; u nähert sich der labilen Ruhelage ($u = \pi$, wenn α positiv, $u = 0$, wenn α negativ) für $t = \pm\infty$, also ohne sie in endlicher Zeit zu erreichen. Doch tritt hier für $e^{\pm t_0} = 0$ der noch engere Spezialfall ein, daß u dauernd in der labilen Ruhelage bleibt.

Da sehr genähert

$$\frac{du}{dt} = m'n' - mn,$$

so kann aus der obigen Kettenbruchentwicklung geschlossen werden, daß $\frac{du}{dt}$ hier in der Tat verschwindet, u periodisch ist, und das rationale Verhältnis $\frac{n'}{n} = \frac{m}{m'}$ sich immer wieder herstellt.

In den Fällen 2. und 3. schwankt $m'\frac{n'}{n} - m$ zwischen den Grenzen $\pm\sqrt{C + 2m|\alpha|}$; im Falle 4. dagegen entweder zwischen $+\sqrt{C \pm 2m|\alpha|}$ oder zwischen $-\sqrt{C \pm 2m|\alpha|}$. Da α äußerst klein ist, so sind auch die Änderungen von n sehr klein.

Die Diskussion der Gleichung 98) stützt sich auf verschiedene Vernachlässigungen und ist daher nicht streng. Der Veränderlichkeit von n^2 ist, immer unter Beiseitelassung von P, unschwer Rechnung zu tragen. Man hat dann die Gleichungen

$$\frac{du}{dt} = m'n' - mn, \qquad \frac{dn}{dt} = n^2\alpha \sin u,$$

zu diskutieren. Setzt man

$$n = \frac{n_0}{1-v} \quad \text{und} \quad m n_0 = m' n',$$

so wird

$$\frac{du}{dt} = -m'n' \frac{v}{1-v}, \qquad \frac{dv}{dt} = n_0 \alpha \sin u.$$

Diese Gleichungen haben die kanonische Form:

$$99) \quad \begin{aligned} \frac{du}{dt} &= \frac{\partial H}{\partial v} \\ \frac{dv}{dt} &= -\frac{\partial H}{\partial u} \end{aligned} \qquad H = n_0 \alpha \cos u + m'n'(v + \log(1-v)).$$

Ein Integral der Gleichungen ist

$$100) \qquad H = c.$$

Die Größe $\frac{du}{dt}$ verschwindet nur für $v = 0$, d. h. nach 100) für $n_0 \alpha \cos u = c$, oder für

$$u = \pm \arccos \frac{c}{n_0 \alpha}.$$

Ist $c > n_0 |\alpha|$, so verschwindet $\frac{du}{dt}$ niemals, entsprechend dem obigen Fall 4. Ist $|c| < n_0 |\alpha|$, so schwankt u periodisch, wie in den Fällen 2. und 3., vorausgesetzt, daß nicht gleichzeitig $\frac{dv}{dt}$ verschwindet. $\frac{du}{dt}$ und $\frac{dv}{dt}$ können nur dann gleichzeitig verschwinden, wenn $c = \pm n_0 |\alpha|$. Für diesen Sonderwert von c treten die Sonderfälle a und c ein: ist $u = 0$ oder π, so bleibt es beständig konstant (Fall a); ist u nicht gleich einem dieser Werte, so tritt der doppelt asymptotische Fall unter c ein.

Diese Untersuchung läßt sich insofern weiter ausdehnen, als sich ohne Schwierigkeit in 97) die Glieder von P berücksichtigen lassen, deren Argumente Vielfache von u sind. Man würde dann Gleichungen der Form 99) erhalten, bei denen

$$H = n_0 \sum \frac{\alpha_i}{i} \cos iu + m'n'(v + \log(1-v)) = c$$

wäre. Die Untersuchung käme dann auf die Feststellung der reellen Wurzeln der Gleichung

$$n_0 \sum \frac{\alpha_i}{i} \cos iu - c = 0$$

heraus. Konvergieren die α_i stark genug, so wird diese nur die Wurzeln $u = 0$ und π haben und die Ergebnisse werden den

obigen wesentlich gleich sein. Daß dieser Fall eintritt, ist bei Pallas zu erwarten, wo m und m' verhältnismäßig große Zahlen sind. Anders kann der Fall liegen, wenn diese Zahlen klein sind, die mittleren Bewegungen sich etwa wie $\frac{m}{m'} = \frac{1}{2}, \frac{1}{3}$ oder $\frac{2}{3}$ verhalten.

Wollte man noch auf weitere Glieder in P und darauf Rücksicht nehmen, daß die Größen $\alpha_{m \cdot m'}$ und $A_{m \cdot m'}$ in den Ausdrücken 91), streng genommen, nicht konstant sind, so werden die Untersuchungen schwieriger[1]).

Nachdem Gauß durch seine zweite Rechnung der allgemeinen Störungen die Zahlenwerte der Koeffizienten α und A gefunden hatte, konnte er die Grenzen, zwischen denen n schwankt, sowie die Länge der Periode, wenn auch nur sehr unsicher, berechnen. Auf dem ersten der oben erwähnten Zettel stehen außer der Zahlenrechnung für die Größen $18L' - 7L$ und C die Gleichungen:

$$„\frac{dd⚴}{dt^2} = 0''.02955 \sin(18♃ - 7⚴ + 198^\circ 39' 8''),$$

$$\frac{dd⚴}{dt^2} = 0''.02955 \sin u \,[= \alpha \sin u],$$

$$\frac{ddu}{dt^2} = -0''.20685 \sin u \,[= -7\alpha \sin u],$$

$$\frac{du^2}{dt^2} = C + 0''.4137 \cos u \,[= C + 14\alpha \cos u],$$

$$dt = \frac{du}{\sqrt{(-0.000\,00159 + 0.000\,00200 \cos u)}}\text{“}.$$

Auf dem zweiten Zettel lauten die Notizen (außer den S. 96 erwähnten Zahlenrechnungen):

$$„\frac{dd⚴}{dt^2} = \alpha \sin(18♃ - 7⚴ + A),$$

[101] $$\frac{ddu}{dt^2} = -7\alpha \sin u, \quad \frac{18d♃}{dt} = 7 + \mu,$$

[102] $$\frac{du^2}{dt^2} = C + 14\alpha \cos u = (14\alpha + C)\cos\psi^2,$$

[103] $$C = \mu\mu - 14\alpha \cos U,$$

$$\cos u = 1 - 2\sin\tfrac{1}{2}u^2,$$

$$\frac{du^2}{dt^2} = C + 14\alpha - 28\alpha \sin\frac{1}{2}u^2 = (C + 14\alpha)\left(1 - \frac{28\alpha}{C + 14\alpha}\sin\frac{1}{2}u^2\right),$$

1) Wir verweisen auf Charlier, Mechanik des Himmels, Leipzig, II. Band, 1907, 11. Abschnitt, § 4—6.

$$\sqrt{\frac{28\alpha}{C+14\alpha}}\cdot\sin\frac{1}{2}u = \sin\psi,$$

$$\sqrt{\frac{7\alpha}{C+14\alpha}}\cdot\cos\frac{1}{2}u.du = \cos\psi.d\psi,$$

$$\frac{du}{dt} = \cos\psi.\sqrt{(C+14\alpha)},$$

$$\frac{d\psi}{dt} = \sqrt{7\alpha}.\sqrt{\left(1-\frac{C+14\alpha}{28\alpha}\sin\psi^2\right)},$$

[104] $$\frac{d\psi}{\sqrt{\left(7\alpha-\frac{C+14\alpha}{4}\sin\psi^2\right)}},\qquad \frac{1}{M(\sqrt{7\alpha},\sqrt{(7\alpha\cos\frac{1}{2}U^2-\frac{1}{4}\mu\mu)})},$$

[105] $$\text{Maximum von } \frac{du}{dt} = \sqrt{\left(28\alpha\sin\frac{1}{2}U^2+\mu\mu\right)}\text{“}.$$

Die Gleichungen 101) und 102) entsprechen den Gleichungen 97) und 98); Gauß schreibt stets kurz dt für $n\,dt$. Nach Gleichung 103) berechnet er C auf dem ersten Zettel; μ und U bedeuten die aus den Beobachtungen festzustellenden Anfangswerte von $\frac{du}{n\,dt} = 18\frac{n'}{n}-7$ und $u = 18L'-7L+A$.

Aus den Werten

$$n' = 299''.12817,$$
$$n = 769''.16512,$$

die Gauß seiner zweiten Rechnung zugrunde legte (Werke VII, 1906, S. 529), ergibt sich

$$\log\mu = 6.2936-10,$$

wo die letzte Stelle schon um einige Einheiten unsicher ist. Der von Gauß angegebene Wert (6.29314) enthält noch eine Stelle mehr und entspräche genau dem Wert $7n = 5384''.15600$ oder $n = 769''.165143$.

Bei der Berechnung von U hat Gauß anscheinend ein Versehen begangen. Es folgt nämlich nach Werke VII, 1906, S. 557 das kritische Glied

$$\frac{du}{n^2dt} = +0''.012\cos(18M'-7M)-0''.027\sin(18M'-7M)$$

oder, indem man beide Teile zusammenzieht:

$$\frac{du}{n^2dt} = 0''.02955\sin(18M'-7M+156^\circ 2' 14'').$$

Der Koeffizient $\alpha = 0''.02955$ stimmt genau mit dem obigem Gauß-

schen Werte überein. Setzt man aber, ebenfalls mit den Werten Werke VII, 1906, S. 529,

$$M' = L' - \pi' = L' - 11^\circ 17' \; 5''.4 \qquad 18\pi' - 7\pi = 75^\circ 5' 16'',$$
$$M = L - \pi = L - 121^\circ \; 8' 54''.5$$

so erhält man

$$\frac{du}{n^2 dt} = 0''.02955 \sin(18L' - 7L + 80^\circ 56' 58'').$$

übrigens in voller Übereinstimmung mit dem von Gauß berechneten Wert im Werke VII, 1906, S. 557 (Glied in $\int n\,dt$). Es ist also $A = 80^\circ 56' 58''$, während Gauß auf den Zetteln mit dem Wert $A = 198^\circ 39' 8''$ gerechnet hat. Zur Berechnung von U setzt er $18L' - 7L = 125^\circ 36' 49''$, ein Wert, der sehr unsicher ist, aber annähernd mit den für verschiedene Epochen nachgerechneten Werten übereinstimmt. Hiermit findet er

$$U = 324^\circ 15' 57'', \qquad \log C = 4.20126\,n - 10,$$

während man mit dem richtigen Wert von A erhält [1]):

$$U = 206^\circ 33' 46'', \qquad \log C = 4.26300 - 10.$$

Es ist aber jedenfalls $C < 14\alpha$ und die Änderung von u periodisch. Die Gleichung 104) gibt die Periode der Libration und die Gleichung 105) gibt die Grenzen, zwischen denen n schwankt. Die Berechnung beider Größen steht auf dem zweiten Zettel neben den Formeln. Man findet die Periode zu 1894 Pallasumläufen = 737 Jupiterumläufen, während Gauß nach seinen Zahlen angibt: 1026,17 Pallasumläufe = 399,07 Jupiterumläufe.

Bei der Berechnung des Ausdruckes 105) ist auf dem Zettel versehentlich $\frac{1}{4}\mu^2$ statt μ^2 gesetzt, womit Gauß die Grenzen von n zu

$$\frac{18}{7} n' \pm 0''.06838.$$

fand (vgl. Werke VII, 1906, S. 559 und 608). Die richtige Rechnung ergibt:

$$\text{für } n\text{: } \frac{18}{7} n' \pm 0''.2153, \qquad \text{für } u\text{: } \pm 156^\circ.$$

und mit den Gaußschen Zahlen:

$$\text{für } n\text{: } \frac{18}{7} n' \pm 0''.0722, \qquad \text{für } u\text{: } \pm 37^\circ 35'.$$

1) In Werke VII, 1906, S. 608, steht versehentlich $\log C = 4.26300$ statt $4.20126\,n$.

Gauß mag die Rechnung auf den Zetteln nur flüchtig gemacht und darauf weniger Wert gelegt haben, da die Grundlagen recht unsicher sind; daher erklärt es sich wohl auch, daß er seine Zahlen nicht geprüft hat. Die Untersuchung gab ihm nur die Unterlagen für seine in der Handschrift „Exposition d'une nouvelle méthode etc.", Werke VII, 1906, S. 468 aufgestellte Behauptung, daß die bei einem rationalen Verhältnis von n und n' auftretenden sekularen Glieder in n bei Berücksichtigung der höheren Potenzen der störenden Masse die periodische Form annehmen. Für die praktische Rechnung empfiehlt er die Beibehaltung des sekularen Gliedes in n, d. h. die Entwicklung des Librationsgliedes nach Potenzen von t, und verfährt hiernach auch selbst (Werke VII, 1906, S. 557).

6. Tafeln für die Jupiterstörungen der Pallas.

Die gesamte Anzahl der Störungsglieder beläuft sich nach S. 60 auf 824 und bei der Berechnung eines einzelnen Pallasortes müßten nach den Ausdrücken in Werke VII, 1906, S. 543—556, alle diese Glieder berechnet werden, wobei schon die Bildung der Argumente die Geduld eines Einzelnen leicht übersteigen würde. Gauß entwarf daher die Werke VII, 1906, S. 572—577 abgedruckten Tafeln, die sich mit den zugehörigen Rechnungen vollständig im Nachlaß vorfanden. Er teilt die periodischen Glieder nach ihren Argumenten in 13 Gruppen und stellt jede Gruppe als Funktion der Differenz $L' - L$ der mittleren Längen von Pallas und Jupiter dar. Mit dieser Größe als Argument können die Zahlen zur Berechnung der Summe einer ganzen solchen Gruppe direkt aus den Tafeln entnommen werden. Die Berechnung der Tafeln, an der sich neben Encke noch Westphal hervorragend beteiligte, wurde in etwas mehr als einem Jahr fertiggestellt.

7. Störungen der Pallas durch Mars.

Von den Marsstörungen hat Gauß nur die Störungen der halben großen Axe, die des ersten Teils $\delta \int n\,dt$ der Epoche und die der Exzentrizität fertig gerechnet und hierbei im wesentlichen wieder dieselbe Methode gebraucht, wie bei den Jupiter- und Saturnstörungen. Die nicht erhebliche Arbeit der Berechnung der übrigen Elemente ist liegen geblieben.

Andererseits finden sich aber Aufzeichnungen im Nachlaß, nach denen Gauß versuchte, eine andere neuartige Methode zur Entwicklung der Komponenten der störenden Kräfte anzuwenden, die sich etwa in folgender Weise beschreiben läßt. Es sind die Ausdrücke

$\frac{x'-x}{\varrho^3}$, $\frac{y'-y}{\varrho^3}$, $\frac{z'-z}{\varrho^3}$ nach Vielfachen der mittleren Anomalien M (Pallas) und M' (Mars) zu entwickeln. Da nicht nur die Exzentrizität der Pallas, sondern auch die des Mars beträchtlich ist, und auch das Verhältnis der beiden halben großen Axen $\frac{a'}{a}$ nicht klein ist, so konvergieren die trigonometrischen Reihen nach allen drei Vielfachen von M, M', $M-M'$ und nach allen drei Parametern $\left(e,\ e' \text{ und } \frac{a'}{a}\right)$ ziemlich schwach und man muß hier auf den Vorteil verzichten, die Entwicklung nach e' an die letzte Stelle zu setzen, wie oben S. 86 bei Gelegenheit der Jupiterstörungen auseinandergesetzt worden ist; vielmehr müßte die Entwicklung nach den Vielfachen aller drei Argumente M, M' und $M'-M$ mit gleichem Recht behandelt werden und dazu reicht die bei den Jupiterstörungen benutzte rein interpolatorische Methode nicht aus, da sie nur die Berücksichtigung zweier Argumente gestattet.

Um diesem Umstand Rechnung zu tragen, scheint Gauß ein gemischtes, teils interpolatorisches, teils analytisches Entwicklungsverfahren angewandt oder wenigstens versucht zu haben. Er teilt den Umkreis für das Argument M in 24 Teile und führt die Entwicklung nach dem Argument M' daher für die Einzelwerte $M = 0, 15^\circ, 30^\circ, \ldots 345^\circ$ aus; bei dieser Entwicklung macht er folgende Betrachtungen.

Für den Abstand der Pallas vom Mars läßt sich die Formel [1])

$$\varrho^2 = A + B\cos(E'-C) + D\cos^2 E'$$

ableiten, wo E' die exzentrische Anomalie des Mars und A, B, C, D Funktionen der exzentrischen, wie auch der mittleren, Anomalie der Pallas sind, also für jeden M-Wert bestimmte Zahlenwerte haben; der Ausdruck für ϱ^2 läßt sich in Faktoren zerlegen:

$$\frac{\varrho^2}{D} = \{\mathrm{M} - \cos(E'-\varphi)\}\{\mathrm{N} - \cos(E'+\varphi)\}.$$

wo M, N, φ aus A, B, C, D zu berechnen sind (siehe Werke VII, 1906, S. 595). Hiermit wird:

$$\frac{x'-x}{\varrho^3} = \frac{x'-x}{D^{\frac{3}{2}}[\mathrm{M}-\cos(E'-\varphi)]^{\frac{3}{2}}[\mathrm{N}-\cos(E'+\varphi)]^{\frac{3}{2}}}.$$

Gauß betrachtet nun die Wurzeln der Gleichung $\varrho^2 = 0$, im be-

1) Die ähnliche Formel $\varrho^2 = (A - a\cos E')^2 + (B - b\sin E')^2 + C^2$ benutzt Gauß in der „Determinatio attractionis etc.“, Werke III, S. 334.

sonderen

$$\mathrm{M} - \cos(E' - \varphi) = 0.$$

Die Wurzeln dieser Gleichung sind, da $\mathrm{M} > 1$ ist,

$$E' = \varphi + k\pi + ix,$$

wo k eine ganze positive oder negative Zahl oder Null bedeutet und wo

$$e^x = \mathrm{M} \pm \sqrt{\mathrm{M}^2 - 1}$$

ist.

Der entsprechende Wert von M' ergibt sich aus der Keplerschen Gleichung

$$M' = E' - e' \sin E' = E' - e' \sin\varphi \cos(E' - \varphi) - e' \cos\varphi \sin(E' - \varphi)$$

oder, wenn man

$$\cos(E' - \varphi) = \mathrm{M} = \frac{1}{\sin\zeta}$$

und damit

$$\sin(E' - \varphi) = i \operatorname{ctg} \zeta$$

setzt:

$$M' = E' - \frac{e' \sin\varphi}{\sin\zeta} - i e' \cos\varphi \operatorname{ctg}\zeta.$$

Setzt man nun

$$\varphi' = \varphi - \frac{e' \sin\varphi}{\sin\zeta}$$

und der Kürze halber

$$y = e' \cos\varphi \operatorname{ctg}\zeta,$$

so wird

$$M' - \varphi' = E' - \varphi - iy,$$

also

$$\begin{aligned}\cos(M' - \varphi') &= \tfrac{1}{2} e^y \{\cos(E' - \varphi) + i \sin(E' - \varphi)\} \\ &\quad + \tfrac{1}{2} e^{-y} \{\cos(E' - \varphi) - i \sin(E' - \varphi)\} \\ &= \tfrac{1}{2} e^y \operatorname{tg} \tfrac{1}{2}\zeta + \tfrac{1}{2} e^{-y} \operatorname{ctg} \tfrac{1}{2}\zeta.\end{aligned}$$

Setzt man noch

$$e^y \operatorname{tg} \tfrac{1}{2}\zeta = \operatorname{tg} \tfrac{1}{2}\zeta',$$

so wird schließlich

$$\cos(M' - \varphi') = \frac{1}{\sin\zeta'}.$$

Die beiden Gleichungen

$$\cos(E' - \varphi) - \mathrm{M} = \cos(E' - \varphi) - \frac{1}{\sin\zeta} = 0$$

und

$$\cos(M'-\varphi') - \frac{1}{\sin\xi'} = 0$$

werden also durch die gleichen Werte von M' befriedigt. Daher hat man

$$\mathrm{M} - \cos(E'-\varphi) = c\left\{\frac{1}{\sin\xi'} - \cos(M'-\varphi')\right\}$$

und, wenn man

$$\lambda = \left\{\frac{1}{\sin\xi'} - \cos(M'-\varphi')\right\}^{\frac{3}{2}}$$

setzt,

$$\frac{x'-x}{\varrho^3} = \frac{x'-x}{c^{\frac{3}{2}} D^{\frac{3}{2}} (\mathrm{N} - \cos(E'+\varphi))^{\frac{3}{2}}} \cdot \frac{1}{\lambda} = \frac{(x'-x)\lambda}{\varrho^3} \cdot \frac{1}{\lambda}.$$

Diese Größe ist also in zwei Faktoren zerlegt und aus dem ersten Faktor $\frac{(x'-x)\lambda}{\varrho^3}$ ist die ungünstige Konvergenz verschwunden, die durch die Größe $(\mathrm{M} - \cos(E'-\varphi))^{-\frac{3}{2}}$ veranlaßt wird. Dieser erste Faktor wird sich also nach der interpolatorischen Methode nach Vielfachen von M' entwickeln lassen, ohne daß man sehr viele Glieder zu berechnen und ohne daß man den Umkreis in allzuviele Teile zu teilen braucht. Dafür wird zwar der zweite Faktor λ mit Berücksichtigung einer größeren Anzahl Glieder entwickelt werden müssen; jedoch läßt dieser sich leicht beliebig weit entwickeln, da es sich hier um nichts anderes als um die von Gauß so vielfach betrachtete Entwicklung des Ausdrucks der Form

$$(\alpha - \beta\cos\psi)^{-\frac{3}{2}}$$

handelt.

Man könnte wohl auch den zweiten Faktor $(\mathrm{N} - \cos(E'+\varphi))^{-\frac{3}{2}}$ in gleicher Weise heraussetzen; doch hat Gauß dieses zweite Verfahren überhaupt nicht weiter verfolgt, weil die Ersparnis an Rechenarbeit doch nicht sehr groß zu sein scheint. Vermutlich hat er den Faktor $(\mathrm{M} - \cos(E'-\varphi))^{-\frac{3}{2}}$ gewählt, weil die Funktion $\mathrm{M} - \cos(E'-\varphi)$ die der reellen Axe am nächsten gelegene Nullstelle in der komplexen M'-Ebene hat. Die entsprechende Wurzel M' hat den kleinsten imaginären Teil; ihr reeller Teil ist φ' und man kann annehmen, daß die Entwicklung nach $M' - \varphi'$ diejenige auf der reellen Axe ist, die am schwächsten konvergiert.

Der Gaußsche Kunstgriff besteht also darin, daß man eine Funktion, deren Entwicklung schwach konvergiert, mit einem

Faktor multipliziert, der die Konvergenz verstärkt. Das Produkt entwickelt man sodann nach der interpolatorischen Methode. Die Entwicklung des reziproken Faktors, mit der nachher wieder zu multiplizieren ist, wird zwar schwach konvergieren; man kann den Faktor aber so wählen, daß seine Entwicklung sich leicht weit genug fortsetzen läßt.

Man kann sich das Verfahren an dem analogen Fall einer Potenzreihenentwicklung vergegenwärtigen:

Es sei die Funktion

$$y = \frac{1}{f(x)}$$

gegeben, wo $f(x)$ eine ganze rationale Funktion sein mag und die Gleichung

$$f(x) = 0$$

keine reelle Wurzel hat. Ihre Wurzeln seien $a_1 \pm i b_1$, $a_2 \pm i b_2$, usw., so daß

$$f(x) = ((x-a_1)^2 + b_1^2)((x-a_2)^2 + b_2^2) \dots .$$

Es wird verlangt, y in einem beliebigen Punkte der reellen Axe so in eine Potenzreihe zu entwickeln, daß diese Reihe möglichst leicht hergestellt werden kann. Wenn b die kleinste der Größen $b_1, b_2, \dots$ ist, so wird die Konvergenz der Entwicklung von y, wenigstens im entsprechenden Punkte a und seiner Umgebung, durch die Nullstelle $a + ib$ begrenzt. Nimmt man den Faktor

$$\lambda = (x-a)^2 + b^2$$

heraus, so konvergiert die Entwicklung von λy stärker als die von y. Es kann also vorteilhaft sein, die Größen λy und $\frac{1}{\lambda}$, jede für sich, zu entwickeln und durch Multiplikation beider Entwicklungen die von y herzustellen. Man wird nämlich bei der Entwicklung von λy weniger hohe Potenzen von x zu berücksichtigen brauchen, wie bei der von y und bei der von $\frac{1}{\lambda}$. Nun ist aber die Entwicklung von

$$\frac{1}{\lambda} = \frac{1}{b^2} \sum_0^\infty (-1)^n \left(\frac{x-a}{b}\right)^{2n}$$

formell leichter herzustellen, als die von y und λy, es wird also durch die Zerlegung ein Vorteil gewonnen.

Solche Erwägungen mögen es gewesen sein, die Gauß hier verfolgt hat. In einer Notiz des Handbuchs 19, Be (S. 76—77),

die Werke VII, 1906, S. 599—600 abgedruckt ist, gibt Gauß eine andere Methode zur Aufsuchung des Faktors λ:

Er bringt die Gleichung $\varrho^2 = 0$ auf die Form:

$$a + b\cos 2E' - \cos(E' + D) = 0$$

und bestimmt ihre Wurzeln E' in der Form $E' = x + y$, wo x reell und y rein imaginär sein soll; die Formeln zur Berechnung von x und $\cos y$ gibt er an; letztere Größe ist reell und größer als Eins. Der entsprechende Wert von M' sei $\xi_0 + i\eta_0$.

Ist λ wieder der zur Entwicklung des Ausdrucks $\frac{1}{\varrho^3}$ aus diesem herauszunehmende Faktor, so ist also $\frac{1}{\lambda}$ in eine Reihe nach Vielfachen von $M' - \xi_0$ zu entwickeln; man wird also λ die Form geben:

$$\lambda = (C - \cos(M' - \xi_0))^{\frac{3}{2}}.$$

Die Konstante C ist so zu bestimmen, daß λ gleichzeitig mit ϱ verschwindet, also

$$C = \cos i\eta_0.$$

Dabei stellt Gauß noch die folgende Betrachtung an:

Sind

$$u = a\cos(nM' + A) \quad \text{und} \quad u' = a'\cos((n+1)M' + A')$$

zwei aufeinanderfolgende Glieder einer konvergenten Reihe, und zwar im besonderen der Entwicklung von $\frac{1}{\lambda}$, so ist:

$$\frac{u'}{u} = \frac{a'}{a}\,\frac{e^{iM' - i(A - A')} + e^{-i(2n+1)M' - i(A + A')}}{1 + e^{-2inM' - 2iA}}.$$

Setzt man

$$M' = \xi + i\eta,$$

so ist

$$\lim_{n \to \infty} \frac{u'}{u} = \lim \frac{a'}{a} e^{-\eta}\{\cos(\xi - A + A') + i\sin(\xi - A + A')\}.$$

Macht man die von Gauß nicht ausdrücklich hervorgehobene Voraussetzung, daß $\lim \frac{u'}{u}$ einen bestimmten Wert hat, so ist dieser höchstens gleich Eins, da die Reihe als konvergent vorausgesetzt ist. Läßt man M' in $M_0' = \xi_0 + i\eta_0$ übergehen und konvergiert die Reihe für alle Werte von M' bis M_0' ausschließlich, so muß $\lim \frac{u'}{u}$ im Punkt M_0' sich beliebig der Eins nähern; es ist

daher:

$$\lim \frac{a'}{a} = e^{\eta_0}, \qquad \lim (A - A') = \xi_0.$$

8. Sekulare Störungen.

Seine Untersuchungen über die sekularen Störungen hat Gauß in der „Determinatio attractionis etc.“ (Werke III, S. 331) veröffentlicht; aus gewissen Äußerlichkeiten kann man vermuten, daß diese Untersuchungen etwa um dieselbe Zeit (1814) ausgeführt worden sind, wie die Berechnung der Marsstörungen der Pallas. Im Handbuch Be S. 82, wo sich Vorarbeiten zur Determinatio finden, fällt nämlich nicht bloß die Ähnlichkeit der Gleichung für ϱ^2 mit der bei der Pallas benutzten auf, sondern dort ist auch ein Zahlenbeispiel von der Pallas entnommen. In der Anzeige der Determinatio Attractionis (Werke III, S. 360) vom Jahre 1818 sagt Gauß, daß er „diese Resultate schon vor vielen Jahren gefunden hat“.

Ebendort (Werke III, S. 357) heißt es: „Vermöge eines, vielleicht bis jetzt noch von niemand ausdrücklich ausgesprochenen, aber aus der physischen Astronomie leicht zu beweisenden Lehrsatzes, sind die Sekularveränderungen einer Planetenbahn durch die Störung eines andern Planeten dieselben, der störende Planet mag eine elliptische Bahn nach Keplers Gesetzen wirklich beschreiben, oder seine Masse mag auf den Umfang der Ellipse in dem Maße verteilt angenommen werden, daß auf Stücke der Ellipse, die sonst in gleich großen Zeiten beschrieben werden, gleich große Anteile an der ganzen Masse kommen: vorausgesetzt, daß die Umlaufszeiten des gestörten und des störenden Planeten nicht in rationalem Verhältnis zu einander stehen“. Man muß hier die Einschränkung machen, daß die Gaußsche Methode nur die ersten Potenzen der störenden Masse berücksichtigt oder daß man nur diese als sekulare Störungen bezeichnet, wie es wohl in der älteren Störungstheorie üblich war. Jedoch wird dadurch z. B. im Falle des Mondes keine genügende Annäherung erreicht; ähnliches tritt ein, wenn die Umlaufszeiten beider Planeten nahezu in einem niedrigzahligen rationalen Verhältnis stehen; in diesem Falle werden die von den dritten Potenzen der störenden Masse abhängenden[1]) Glieder, die die Sekularbeschleunigung hauptsächlich bedingen, merklich.

---

1) Vgl. Brendel, Theorie der kleinen Planeten, 2. Teil, S. 49—51 u. sonst (Abhandlungen der K. Gesellschaft der Wissenschaften zu Göttingen, Math.-phys. Klasse, Neue Folge, Bd. VI, Nr. 4, Berlin 1909).

9. Schlußbemerkung.

Mit den Jahren 1816—17 schließen die Arbeiten Gauß' aus dem Gebiete der theoretischen Astronomie gänzlich ab. Seine späteren astronomischen Untersuchungen gehören in das Gebiet der beobachtenden und sphärischen Astronomie, während sich der Schwerpunkt seiner Arbeiten überhaupt auf verwandte Gebiete verschob; in den nächstfolgenden Jahren steht besonders die Geodäsie im Vordergrund seiner Tätigkeit.
